

1. [Preface](#)
2. Unit 1. The Chemistry of Life
  1. The Study of Life
    1. [Introduction](#)
    2. [The Science of Biology](#)
    3. [Themes and Concepts of Biology](#)
  2. Biological Macromolecules
    1. [Introduction](#)
    2. [Synthesis of Biological Macromolecules](#)
    3. [Carbohydrates](#)
    4. [Lipids](#)
    5. [Proteins](#)
    6. [Nucleic Acids](#)
3. Unit 2. The Cell
  1. Cell Structure
    1. [Introduction](#)
    2. [Studying Cells](#)
    3. [Prokaryotic Cells](#)
    4. [Eukaryotic Cells](#)
    5. [The Endomembrane System and Proteins](#)
    6. [The Cytoskeleton](#)
    7. [Connections between Cells and Cellular Activities](#)
  2. Cell Reproduction
    1. [Introduction](#)
    2. [Cell Division](#)
    3. [The Cell Cycle](#)
    4. [Control of the Cell Cycle](#)
    5. [Prokaryotic Cell Division](#)
4. Unit 3. Genetics
  1. Meiosis and Sexual Reproduction
    1. [Introduction](#)

2. [The Process of Meiosis](#)
  3. [Sexual Reproduction](#)
2. Mendel's Experiments and Heredity
  1. [Introduction](#)
  2. [Mendel's Experiments and the Laws of Probability](#)
  3. [Characteristics and Traits](#)
  4. [Laws of Inheritance](#)
3. Modern Understandings of Inheritance
  1. [Introduction](#)
  2. [Chromosomal Theory and Genetic Linkage](#)
  3. [Chromosomal Basis of Inherited Disorders](#)
4. DNA Structure and Function
  1. [Introduction](#)
  2. [Historical Basis of Modern Understanding](#)
  3. [DNA Structure and Sequencing](#)
  4. [Basics of DNA Replication](#)
  5. [DNA Replication in Prokaryotes](#)
  6. [DNA Replication in Eukaryotes](#)
  7. [DNA Repair](#)
5. Genes and Proteins
  1. [Introduction](#)
  2. [The Genetic Code](#)
  3. [Prokaryotic Transcription](#)
  4. [Eukaryotic Transcription](#)
  5. [RNA Processing in Eukaryotes](#)
  6. [Ribosomes and Protein Synthesis](#)
6. Gene Expression
  1. [Introduction](#)
  2. [Regulation of Gene Expression](#)
  3. [Prokaryotic Gene Regulation](#)
  4. [Eukaryotic Epigenetic Gene Regulation](#)
  5. [Eukaryotic Transcription Gene Regulation](#)



6. [Eukaryotic Post-transcriptional Gene Regulation](#)
7. [Eukaryotic Translational and Post-translational Gene Regulation](#)
8. [Cancer and Gene Regulation](#)
7. Biotechnology and Genomics
  1. [Introduction](#)
  2. [Biotechnology](#)
  3. [Mapping Genomes](#)
  4. [Whole-Genome Sequencing](#)
  5. [Applying Genomics](#)
  6. [Genomics and Proteomics](#)
5. Unit 4. Evolutionary Processes
  1. Evolution and the Origin of Species
    1. [Introduction](#)
    2. [Understanding Evolution](#)
    3. [Formation of New Species](#)
    4. [Reconnection and Rates of Speciation](#)
  2. The Evolution of Populations
    1. [Introduction](#)
    2. [Population Evolution](#)
    3. [Population Genetics](#)
    4. [Adaptive Evolution](#)
  3. Phylogenies and the History of Life
    1. [Introduction](#)
    2. [Organizing Life on Earth](#)
    3. [Determining Evolutionary Relationships](#)
    4. [Perspectives on the Phylogenetic Tree](#)
6. [The Periodic Table of Elements](#)
7. [Geological Time](#)
8. [Measurements and the Metric System](#)

## Preface

Biology is designed for multi-semester biology courses for science majors. It is grounded on an evolutionary basis and includes exciting features that highlight careers in the biological sciences and everyday applications of the concepts at hand. To meet the needs of today's instructors and students, some content has been strategically condensed while maintaining the overall scope and coverage of traditional texts for this course. Instructors can customize the book, adapting it to the approach that works best in their classroom. Biology also includes an innovative art program that incorporates critical thinking and clicker questions to help students understand—and apply—key concepts.

Welcome to *Biology*, an OpenStax resource. This textbook was written to increase student access to high-quality learning materials, maintaining highest standards of academic rigor at little to no cost.

## About OpenStax

OpenStax is a nonprofit based at Rice University, and it's our mission to improve student access to education. Our first openly licensed college textbook was published in 2012, and our library has since scaled to over 20 books for college and AP courses used by hundreds of thousands of students. Our adaptive learning technology, designed to improve learning outcomes through personalized educational paths, is being piloted in college courses throughout the country. Through our partnerships with philanthropic foundations and our alliance with other educational resource organizations, OpenStax is breaking down the most common barriers to learning and empowering students and instructors to succeed.

## About OpenStax's Resources

### Customization

*Biology* is licensed under a Creative Commons Attribution 4.0 International (CC BY) license, which means that you can distribute, remix, and build

upon the content, as long as you provide attribution to OpenStax and its content contributors.

Because our books are openly licensed, you are free to use the entire book or pick and choose the sections that are most relevant to the needs of your course. Feel free to remix the content by assigning your students certain chapters and sections in your syllabus, in the order that you prefer. You can even provide a direct link in your syllabus to the sections in the web view of your book.

Instructors also have the option of creating a customized version of their OpenStax book. The custom version can be made available to students in low-cost print or digital form through their campus bookstore. Visit your book page on [openstax.org](https://openstax.org) for more information.

## **Errata**

All OpenStax textbooks undergo a rigorous review process. However, like any professional-grade textbook, errors sometimes occur. Since our books are web based, we can make updates periodically when deemed pedagogically necessary. If you have a correction to suggest, submit it through the link on your book page on [openstax.org](https://openstax.org). Subject matter experts review all errata suggestions. OpenStax is committed to remaining transparent about all updates, so you will also find a list of past errata changes on your book page on [openstax.org](https://openstax.org).

## **Format**

You can access this textbook for free in web view or PDF through [openstax.org](https://openstax.org), and in low-cost print and iBooks editions.

## **About Biology**

*Biology* is designed to cover the scope and sequence requirements of a typical two-semester biology course for science majors. The text provides comprehensive coverage of foundational research and core biology concepts through an evolutionary lens. *Biology* includes rich features that engage students in scientific inquiry, highlight careers in the biological sciences, and offer everyday applications. The book also includes clicker questions to help students understand—and apply—key concepts.

## Coverage and Scope

In developing *Biology*, we listened to hundreds of General Biology instructors who readily provided feedback about their courses, students, challenges, and hopes for innovation. The expense of textbooks and related items did prove to be a barrier to learning. But more importantly, these teachers suggested improvements for the textbook, which would ultimately lead to more meaningful and memorable learning experiences for students.

The result is a book that addresses a core organizational reality of the course and its materials—the sheer breadth of the topical coverage. We provide a thorough treatment of biology’s foundational concepts while condensing selected topics in response to the market’s request for a textbook with a scope that is manageable for instructors and students alike. We also strive to make biology, as a discipline, interesting and accessible to students. In addition to a comprehensive coverage of core concepts and foundational research, we have incorporated features that draw learners into the discipline in meaningful ways.

The pedagogical choices, chapter arrangements, and learning objective fulfillment were developed and vetted with the feedback of another one hundred reviewers, who thoroughly read the material and offered detailed critical commentary.

**Unit 1: The Chemistry of Life.** Our opening unit introduces students to the sciences, including the scientific method and the fundamental concepts of chemistry and physics that provide a framework within which learners comprehend biological processes.

Unit 2: **The Cell.** Students will gain solid understanding of the structures, functions, and processes of the most basic unit of life: the cell.

Unit 3: **Genetics.** Our comprehensive genetics unit takes learners from the earliest experiments that revealed the basis of genetics through the intricacies of DNA to current applications in the emerging studies of biotechnology and genomics.

Unit 4: **Evolutionary Processes.** The core concepts of evolution are discussed in this unit with examples illustrating evolutionary processes. Additionally, the evolutionary basis of biology reappears throughout the textbook in general discussion and is reinforced through special call-out features highlighting specific evolution-based topics.

Unit 5: **Biological Diversity.** The diversity of life is explored with detailed study of various organisms and discussion of emerging phylogenetic relationships. This unit moves from viruses to living organisms like bacteria, discusses the organisms formerly grouped as protists, and devotes multiple chapters to plant and animal life.

Unit 6: **Plant Structure and Function.** Our plant unit thoroughly covers the fundamental knowledge of plant life essential to an introductory biology course.

Unit 7: **Animal Structure and Function.** An introduction to the form and function of the animal body is followed by chapters on specific body systems and processes. This unit touches on the biology of all organisms while maintaining an engaging focus on human anatomy and physiology that helps students connect to the topics.

Unit 8: **Ecology.** Ecological concepts are broadly covered in this unit, with features highlighting localized, real-world issues of conservation and biodiversity.

## **Pedagogical Foundation and Features**

*Biology* is grounded in a solid scientific base, with features that engage the students in scientific inquiry, including:

**Evolution Connection** features uphold the importance of evolution to all biological study through discussions like “The Evolution of Metabolic Pathways” and “Algae and Evolutionary Paths to Photosynthesis.”

**Scientific Method Connection** call-outs walk students through actual or thought experiments that elucidate the steps of the scientific process as applied to the topic. Features include “Determining the Time Spent in Cell Cycle Stages” and “Testing the Hypothesis of Independent Assortment.”

**Career Connection** features present information on a variety of careers in the biological sciences, introducing students to the educational requirements and day-to-day work life of a variety of professions, such as microbiologist, ecologist, neurologist, and forensic scientist.

**Everyday Connection** features tie biological concepts to emerging issues and discuss science in terms of everyday life. Topics include “Chesapeake Bay” and “Can Snail Venom Be Used as a Pharmacological Pain Killer?”

## **Art and Animations That Engage**

Our art program takes a straightforward approach designed to help students learn the concepts of biology through simple, effective illustrations, photos, and micrographs. *Biology* also incorporates links to relevant animations and interactive exercises that help bring biology to life for students.

**Art Connection** features call out core figures in each chapter for student study. Questions about key figures, including clicker questions that can be used in the classroom, engage students’ critical thinking to ensure genuine understanding.

**Link to Learning** features direct students to online interactive exercises and animations to add a fuller context to core content.

## **Additional Resources**

## **Student and Instructor Resources**

We've compiled additional resources for both students and instructors, including Getting Started Guides, an instructor solution manual, supplemental test items, and PowerPoint slides. Instructor resources require a verified instructor account, which can be requested on your [openstax.org](https://openstax.org) log-in. Take advantage of these resources to supplement your OpenStax book.

## **Partner Resources**

OpenStax Partners are our allies in the mission to make high-quality learning materials affordable and accessible to students and instructors everywhere. Their tools integrate seamlessly with our OpenStax titles at a low cost. To access the partner resources for your text, visit your book page on [openstax.org](https://openstax.org).

## **About the Authors**

### **Senior Contributing Authors**

Yael Avissar (Cell Biology), Rhode Island College

Jung Choi (Genetics), Georgia Institute of Technology

Jean DeSaix (Evolution), University of North Carolina at Chapel Hill

Vladimir Jurukovski (Animal Physiology), Suffolk County Community College

Robert Wise (Plant Biology), University of Wisconsin, Oshkosh

Connie Rye (General Content Lead), East Mississippi Community College

### **Contributing Authors and Reviewers**

Julie Adams, Aurora University

Summer Allen, Brown University

James Bader, Case Western Reserve University

David Bailey, St. Norbert College

Mark Belk, Brigham Young University

Nancy Boury, Iowa State University

Lisa Bonneau, Metropolitan Community College – Blue River  
Graciela Brelles-Marino, California State University Pomona  
Mark Browning, Purdue University  
Sue Chaplin, University of St. Thomas  
George Cline, Jacksonville State University  
Deb Cook, Georgia Gwinnett College  
Diane Day, Clayton State University  
Frank Dirrigl, The University of Texas Pan American  
Waneene Dorsey, Grambling State University  
Nick Downey, University of Wisconsin La Crosse  
Rick Duhrkopf, Baylor University  
Kristy Duran, Adams State University  
Stan Eisen, Christian Brothers University  
Brent Ewers, University of Wyoming  
Myriam Feldman, Lake Washington Institute of Technology  
Michael Fine, Virginia Commonwealth University  
Linda Flora, Delaware County Community College  
Thomas Freeland, Walsh University  
David Gris , Texas A & M University – Corpus Christi  
Andrea Hazard, SUNY Cortland  
Michael Hedrick, University of North Texas  
Linda Hensel, Mercer University  
Mark Kopeny, University of Virginia  
Norman Johnson, University of Massachusetts Amherst  
Grace Lasker, Lake Washington Institute of Technology; Walden University  
Sandy Latourelle, SUNY Plattsburgh  
Theo Light, Shippensburg University  
Clark Lindgren, Grinnell College  
James Malcolm, University of Redlands  
Mark Meade, Jacksonville State University  
Richard Merritt, Houston Community College  
James Mickle, North Carolina State University  
Jasleen Mishra, Houston Community College  
Dudley Moon, Albany College of Pharmacy and Health Sciences  
Shobhana Natarajan, Brookhaven College  
Jonas Okeagu, Fayetteville State University  
Diana Oliveras, University of Colorado Boulder



John Peters, College of Charleston  
Joel Piperberg, Millersville University  
Johanna Porter-Kelley, Winston-Salem State University  
Robyn Puffenbarger, Bridgewater College  
Dennis Revie, California Lutheran University  
Ann Rushing, Baylor University  
Sangha Saha, City College of Chicago  
Edward Saiff, Ramapo College of New Jersey  
Brian Shmaefsky, Lone Star College System  
Robert Sizemore, Alcorn State University  
Marc Smith, Sinclair Community College  
Frederick Spiegel, University of Arkansas  
Frederick Sproull, La Roche College  
Bob Sullivan, Marist College  
Mark Sutherland, Hendrix College  
Toure Thompson, Alabama A&M University  
Scott Thomson, University of Wisconsin – Parkside  
Allison van de Meene, University of Melbourne  
Mary White, Southeastern Louisiana University  
Steven Wilt, Bellarmine University  
James Wise, Hampton University  
Renna Wolfe  
Virginia Young, Mercer University  
Leslie Zeman, University of Washington  
Daniel Zurek, Pittsburg State University  
Shobhana Natarajan, Alcon Laboratories, Inc.

## Introduction

class="introduction"

This NASA image is a composite of several satellite-based views of Earth. To make the whole-Earth image, NASA scientists combine observations of different parts of the planet. (credit: NASA/GSFC/NOAA/USGS)

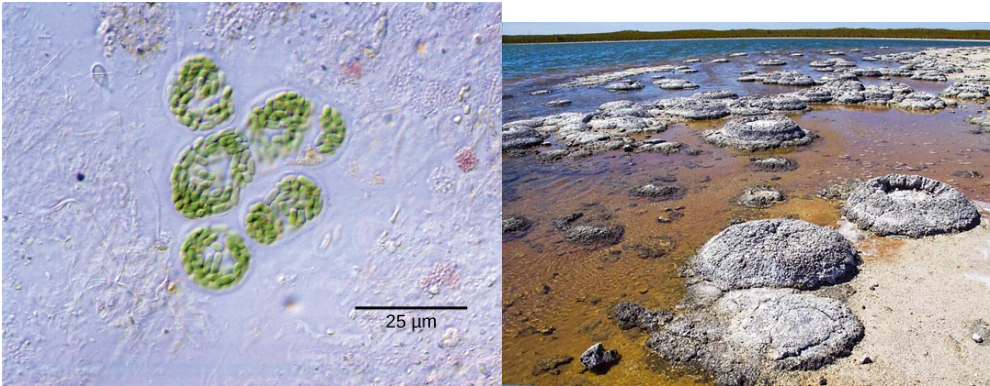


Viewed from space, Earth offers no clues about the diversity of life forms that reside there. The first forms of life on Earth are thought to have been microorganisms that existed for billions of years in the ocean before plants and animals appeared. The mammals, birds, and flowers so familiar to us are all relatively recent, originating 130 to 200 million years ago. Humans have inhabited this planet for only the last 2.5 million years, and only in the last 200,000 years have humans started looking like we do today.

## The Science of Biology

By the end of this section, you will be able to:

- Identify the shared characteristics of the natural sciences
- Summarize the steps of the scientific method
- Compare inductive reasoning with deductive reasoning
- Describe the goals of basic science and applied science



Formerly called blue-green algae, these (a) cyanobacteria, shown here at 300x magnification under a light microscope, are some of Earth's oldest life forms.

These (b) stromatolites along the shores of Lake Thetis in Western Australia are ancient structures formed by the layering of cyanobacteria in shallow waters. (credit a: modification of work by NASA; credit b: modification of work by Ruth Ellison; scale-bar data from Matt Russell)

What is biology? In simple terms, **biology** is the study of living organisms and their interactions with one another and their environments. This is a very broad definition because the scope of biology is vast. Biologists may study anything from the microscopic or submicroscopic view of a cell to ecosystems and the whole living planet ([\[link\]](#)). Listening to the daily news, you will quickly realize how many aspects of biology are discussed every day. For example, recent news topics include *Escherichia coli* ([\[link\]](#)) outbreaks in spinach and *Salmonella* contamination in peanut butter. Other subjects include efforts toward finding a cure for AIDS, Alzheimer's

disease, and cancer. On a global scale, many researchers are committed to finding ways to protect the planet, solve environmental issues, and reduce the effects of climate change. All of these diverse endeavors are related to different facets of the discipline of biology.



*Escherichia coli* (*E. coli*) bacteria, seen in this scanning electron micrograph, are normal residents of our digestive tracts that aid in the absorption of vitamin K and other nutrients. However, virulent strains are sometimes responsible for disease outbreaks. (credit: Eric Erbe, digital colorization by Christopher Pooley, both of USDA, ARS, EMU)

## The Process of Science

Biology is a science, but what exactly is science? What does the study of biology share with other scientific disciplines? **Science** (from the Latin *scientia*, meaning “knowledge”) can be defined as knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method. It becomes clear from this definition

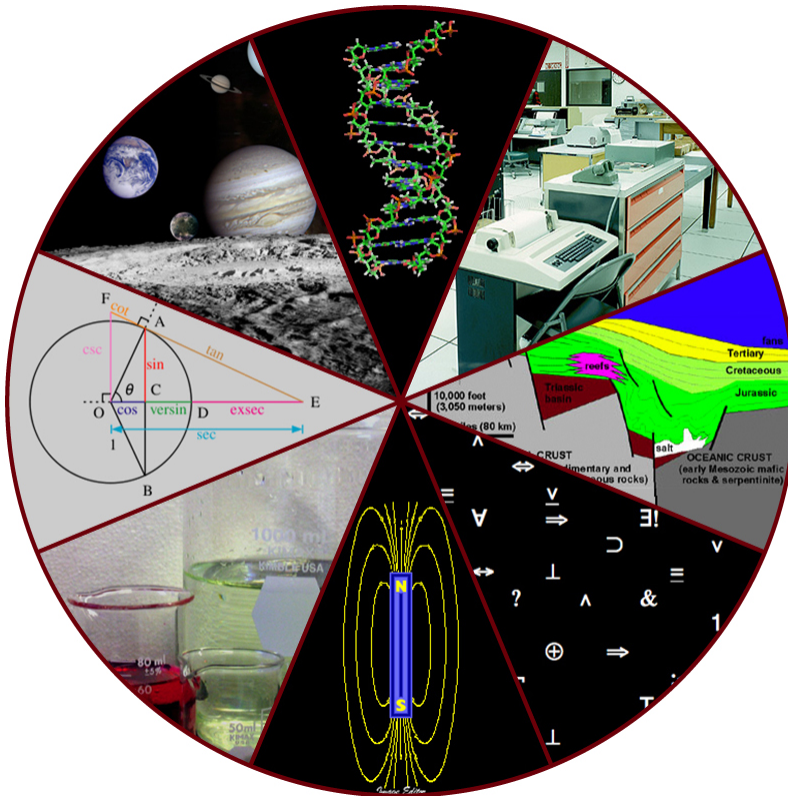
that the application of the scientific method plays a major role in science. The **scientific method** is a method of research with defined steps that include experiments and careful observation.

The steps of the scientific method will be examined in detail later, but one of the most important aspects of this method is the testing of hypotheses by means of repeatable experiments. A **hypothesis** is a suggested explanation for an event, which can be tested. Although using the scientific method is inherent to science, it is inadequate in determining what science is. This is because it is relatively easy to apply the scientific method to disciplines such as physics and chemistry, but when it comes to disciplines like archaeology, psychology, and geology, the scientific method becomes less applicable as it becomes more difficult to repeat experiments.

These areas of study are still sciences, however. Consider archeology—even though one cannot perform repeatable experiments, hypotheses may still be supported. For instance, an archeologist can hypothesize that an ancient culture existed based on finding a piece of pottery. Further hypotheses could be made about various characteristics of this culture, and these hypotheses may be found to be correct or false through continued support or contradictions from other findings. A hypothesis may become a verified theory. A **theory** is a tested and confirmed explanation for observations or phenomena. Science may be better defined as fields of study that attempt to comprehend the nature of the universe.

## Natural Sciences

What would you expect to see in a museum of natural sciences? Frogs? Plants? Dinosaur skeletons? Exhibits about how the brain functions? A planetarium? Gems and minerals? Or, maybe all of the above? Science includes such diverse fields as astronomy, biology, computer sciences, geology, logic, physics, chemistry, and mathematics ([\[link\]](#)). However, those fields of science related to the physical world and its phenomena and processes are considered **natural sciences**. Thus, a museum of natural sciences might contain any of the items listed above.



The diversity of scientific fields includes astronomy, biology, computer science, geology, logic, physics, chemistry, mathematics, and many other fields. (credit: “Image Editor”/Flickr)

There is no complete agreement when it comes to defining what the natural sciences include, however. For some experts, the natural sciences are astronomy, biology, chemistry, earth science, and physics. Other scholars choose to divide natural sciences into **life sciences**, which study living things and include biology, and **physical sciences**, which study nonliving matter and include astronomy, geology, physics, and chemistry. Some disciplines such as biophysics and biochemistry build on both life and physical sciences and are interdisciplinary. Natural sciences are sometimes referred to as “hard science” because they rely on the use of quantitative data; social sciences that study society and human behavior are more likely to use qualitative assessments to drive investigations and findings.



Not surprisingly, the natural science of biology has many branches or subdisciplines. Cell biologists study cell structure and function, while biologists who study anatomy investigate the structure of an entire organism. Those biologists studying physiology, however, focus on the internal functioning of an organism. Some areas of biology focus on only particular types of living things. For example, botanists explore plants, while zoologists specialize in animals.

## **Scientific Reasoning**

One thing is common to all forms of science: an ultimate goal “to know.” Curiosity and inquiry are the driving forces for the development of science. Scientists seek to understand the world and the way it operates. To do this, they use two methods of logical thinking: inductive reasoning and deductive reasoning.

**Inductive reasoning** is a form of logical thinking that uses related observations to arrive at a general conclusion. This type of reasoning is common in descriptive science. A life scientist such as a biologist makes observations and records them. These data can be qualitative or quantitative, and the raw data can be supplemented with drawings, pictures, photos, or videos. From many observations, the scientist can infer conclusions (inductions) based on evidence. Inductive reasoning involves formulating generalizations inferred from careful observation and the analysis of a large amount of data. Brain studies provide an example. In this type of research, many live brains are observed while people are doing a specific activity, such as viewing images of food. The part of the brain that “lights up” during this activity is then predicted to be the part controlling the response to the selected stimulus, in this case, images of food. The “lighting up” of the various areas of the brain is caused by excess absorption of radioactive sugar derivatives by active areas of the brain. The resultant increase in radioactivity is observed by a scanner. Then, researchers can stimulate that part of the brain to see if similar responses result.

Deductive reasoning or deduction is the type of logic used in hypothesis-based science. In deductive reason, the pattern of thinking moves in the opposite direction as compared to inductive reasoning. **Deductive reasoning** is a form of logical thinking that uses a general principle or law to forecast specific results. From those general principles, a scientist can extrapolate and predict the specific results that would be valid as long as the general principles are valid. Studies in climate change can illustrate this type of reasoning. For example, scientists may predict that if the climate becomes warmer in a particular region, then the distribution of plants and animals should change. These predictions have been made and tested, and many such changes have been found, such as the modification of arable areas for agriculture, with change based on temperature averages.

Both types of logical thinking are related to the two main pathways of scientific study: descriptive science and hypothesis-based science. **Descriptive (or discovery) science**, which is usually inductive, aims to observe, explore, and discover, while **hypothesis-based science**, which is usually deductive, begins with a specific question or problem and a potential answer or solution that can be tested. The boundary between these two forms of study is often blurred, and most scientific endeavors combine both approaches. The fuzzy boundary becomes apparent when thinking about how easily observation can lead to specific questions. For example, a gentleman in the 1940s observed that the burr seeds that stuck to his clothes and his dog's fur had a tiny hook structure. On closer inspection, he discovered that the burrs' gripping device was more reliable than a zipper. He eventually developed a company and produced the hook-and-loop fastener popularly known today as Velcro. Descriptive science and hypothesis-based science are in continuous dialogue.

## The Scientific Method

Biologists study the living world by posing questions about it and seeking science-based responses. This approach is common to other sciences as well and is often referred to as the scientific method. The scientific method was used even in ancient times, but it was first documented by England's Sir Francis Bacon (1561–1626) ([\[link\]](#)), who set up inductive methods for scientific inquiry. The scientific method is not exclusively used by



biologists but can be applied to almost all fields of study as a logical, rational problem-solving method.



Sir Francis Bacon  
(1561–1626) is  
credited with being  
the first to define  
the scientific  
method. (credit:  
Paul van Somer)

The scientific process typically starts with an observation (often a problem to be solved) that leads to a question. Let's think about a simple problem that starts with an observation and apply the scientific method to solve the problem. One Monday morning, a student arrives at class and quickly discovers that the classroom is too warm. That is an observation that also describes a problem: the classroom is too warm. The student then asks a question: "Why is the classroom so warm?"

## Proposing a Hypothesis

Recall that a hypothesis is a suggested explanation that can be tested. To solve a problem, several hypotheses may be proposed. For example, one hypothesis might be, “The classroom is warm because no one turned on the air conditioning.” But there could be other responses to the question, and therefore other hypotheses may be proposed. A second hypothesis might be, “The classroom is warm because there is a power failure, and so the air conditioning doesn’t work.”

Once a hypothesis has been selected, the student can make a prediction. A prediction is similar to a hypothesis but it typically has the format “If . . . then . . . .” For example, the prediction for the first hypothesis might be, “*If* the student turns on the air conditioning, *then* the classroom will no longer be too warm.”

## Testing a Hypothesis

A valid hypothesis must be testable. It should also be **falsifiable**, meaning that it can be disproven by experimental results. Importantly, science does not claim to “prove” anything because scientific understandings are always subject to modification with further information. This step—openness to disproving ideas—is what distinguishes sciences from non-sciences. The presence of the supernatural, for instance, is neither testable nor falsifiable. To test a hypothesis, a researcher will conduct one or more experiments designed to eliminate one or more of the hypotheses. Each experiment will have one or more variables and one or more controls. A **variable** is any part of the experiment that can vary or change during the experiment. The **control group** contains every feature of the experimental group except it is not given the manipulation that is hypothesized about. Therefore, if the results of the experimental group differ from the control group, the difference must be due to the hypothesized manipulation, rather than some outside factor. Look for the variables and controls in the examples that follow. To test the first hypothesis, the student would find out if the air conditioning is on. If the air conditioning is turned on but does not work, there should be another reason, and this hypothesis should be rejected. To

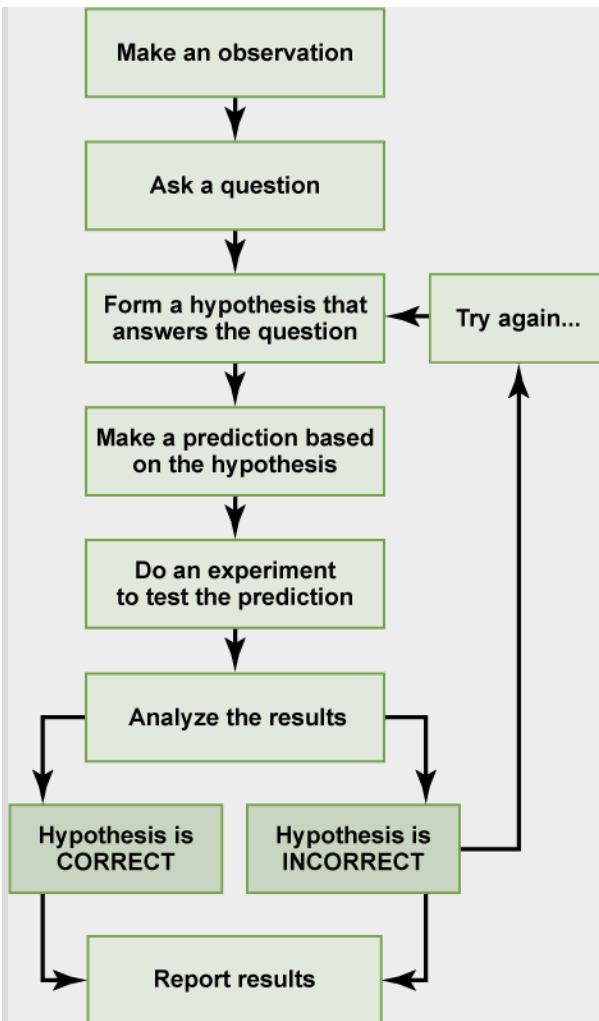
test the second hypothesis, the student could check if the lights in the classroom are functional. If so, there is no power failure and this hypothesis should be rejected. Each hypothesis should be tested by carrying out appropriate experiments. Be aware that rejecting one hypothesis does not determine whether or not the other hypotheses can be accepted; it simply eliminates one hypothesis that is not valid ([\[link\]](#)). Using the scientific method, the hypotheses that are inconsistent with experimental data are rejected.

While this “warm classroom” example is based on observational results, other hypotheses and experiments might have clearer controls. For instance, a student might attend class on Monday and realize she had difficulty concentrating on the lecture. One observation to explain this occurrence might be, “When I eat breakfast before class, I am better able to pay attention.” The student could then design an experiment with a control to test this hypothesis.

In hypothesis-based science, specific results are predicted from a general premise. This type of reasoning is called deductive reasoning: deduction proceeds from the general to the particular. But the reverse of the process is also possible: sometimes, scientists reach a general conclusion from a number of specific observations. This type of reasoning is called inductive reasoning, and it proceeds from the particular to the general. Inductive and deductive reasoning are often used in tandem to advance scientific knowledge ([\[link\]](#)). In recent years a new approach of testing hypotheses has developed as a result of an exponential growth of data deposited in various databases. Using computer algorithms and statistical analyses of data in databases, a new field of so-called “data research” (also referred to as “in silico” research) provides new methods of data analyses and their interpretation. This will increase the demand for specialists in both biology and computer science, a promising career opportunity.

**Note:**

Art Connection



The scientific method consists of a series of well-defined steps. If a hypothesis is not supported by experimental data, a new hypothesis can be proposed.

In the example below, the scientific method is used to solve an everyday problem. Order the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items). Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation

2. Question
3. Hypothesis (answer)
4. Prediction
5. Experiment
6. Result

- a. There is something wrong with the electrical outlet.
- b. If something is wrong with the outlet, my coffeemaker also won't work when plugged into it.
- c. My toaster doesn't toast my bread.
- d. I plug my coffee maker into the outlet.
- e. My coffeemaker works.
- f. Why doesn't my toaster work?

## Note:

### Art Connection

Two Types of Reasoning	
<b>Inductive reasoning:</b> from a number of observations, a general conclusion is drawn.	<b>Deductive reasoning:</b> from a general premise, specific results are predicted.
<b>Observations</b>	<b>General premise</b>
<ul style="list-style-type: none"> <li>Members of a species are not all the same.</li> <li>Individuals compete for resources.</li> <li>Species are generally adapted to their environment.</li> </ul>	Individuals most adapted to their environment are more likely to survive and pass their traits on to the next generation.
↓	↓
<b>Conclusion</b>	<b>Predicted results</b>
Individuals most adapted to their environment are more likely to survive and pass their traits to the next generation.	If the average temperature in an ecosystem increases due to climate change, individuals better adapted to warmer temperatures will outcompete those that are not.

Scientists use two types of reasoning, inductive and deductive reasoning, to advance scientific knowledge. As is the case in this example, the conclusion from inductive reasoning can often become the premise for inductive reasoning.

Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.
2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA, separate into daughter cells during cell division. Therefore, DNA is the genetic material.
4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

The scientific method may seem too rigid and structured. It is important to keep in mind that, although scientists often follow this sequence, there is flexibility. Sometimes an experiment leads to conclusions that favor a change in approach; often, an experiment brings entirely new scientific questions to the puzzle. Many times, science does not operate in a linear fashion; instead, scientists continually draw inferences and make generalizations, finding patterns as their research proceeds. Scientific reasoning is more complex than the scientific method alone suggests. Notice, too, that the scientific method can be applied to solving problems that aren't necessarily scientific in nature.

## Two Types of Science: Basic Science and Applied Science

The scientific community has been debating for the last few decades about the value of different types of science. Is it valuable to pursue science for the sake of simply gaining knowledge, or does scientific knowledge only have worth if we can apply it to solving a specific problem or to bettering our lives? This question focuses on the differences between two types of science: basic science and applied science.

**Basic science** or “pure” science seeks to expand knowledge regardless of the short-term application of that knowledge. It is not focused on developing a product or a service of immediate public or commercial value. The immediate goal of basic science is knowledge for knowledge’s sake, though this does not mean that, in the end, it may not result in a practical application.

In contrast, **applied science** or “technology,” aims to use science to solve real-world problems, making it possible, for example, to improve a crop yield, find a cure for a particular disease, or save animals threatened by a natural disaster ([\[link\]](#)). In applied science, the problem is usually defined for the researcher.



After Hurricane Ike struck the Gulf Coast in 2008, the U.S. Fish and Wildlife Service rescued this

brown pelican. Thanks to applied science, scientists knew how to rehabilitate the bird. (credit: FEMA)

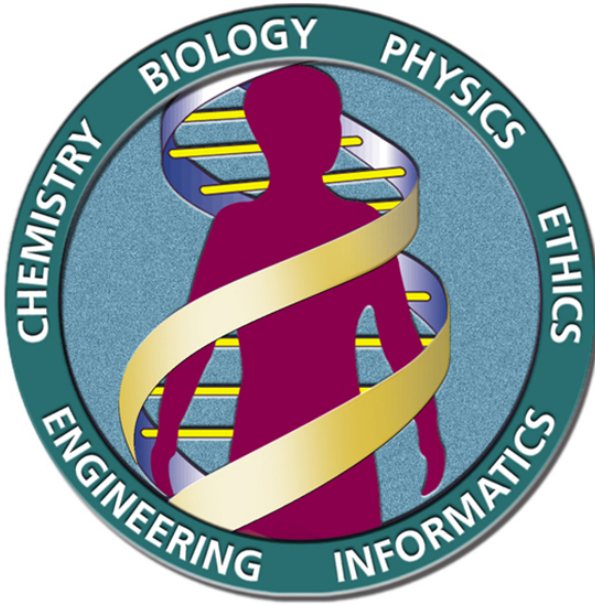
Some individuals may perceive applied science as “useful” and basic science as “useless.” A question these people might pose to a scientist advocating knowledge acquisition would be, “What for?” A careful look at the history of science, however, reveals that basic knowledge has resulted in many remarkable applications of great value. Many scientists think that a basic understanding of science is necessary before an application is developed; therefore, applied science relies on the results generated through basic science. Other scientists think that it is time to move on from basic science and instead to find solutions to actual problems. Both approaches are valid. It is true that there are problems that demand immediate attention; however, few solutions would be found without the help of the wide knowledge foundation generated through basic science.

One example of how basic and applied science can work together to solve practical problems occurred after the discovery of DNA structure led to an understanding of the molecular mechanisms governing DNA replication. Strands of DNA, unique in every human, are found in our cells, where they provide the instructions necessary for life. During DNA replication, DNA makes new copies of itself, shortly before a cell divides. Understanding the mechanisms of DNA replication enabled scientists to develop laboratory techniques that are now used to identify genetic diseases, pinpoint individuals who were at a crime scene, and determine paternity. Without basic science, it is unlikely that applied science would exist.

Another example of the link between basic and applied research is the Human Genome Project, a study in which each human chromosome was analyzed and mapped to determine the precise sequence of DNA subunits and the exact location of each gene. (The gene is the basic unit of heredity; an individual’s complete collection of genes is his or her genome.) Other less complex organisms have also been studied as part of this project in



order to gain a better understanding of human chromosomes. The Human Genome Project ([\[link\]](#)) relied on basic research carried out with simple organisms and, later, with the human genome. An important end goal eventually became using the data for applied research, seeking cures and early diagnoses for genetically related diseases.



The Human Genome Project was a 13-year collaborative effort among researchers working in several different fields of science. The project, which sequenced the entire human genome, was completed in 2003. (credit: the U.S. Department of Energy Genome Programs (<http://genomics.energy.gov>))

While research efforts in both basic science and applied science are usually carefully planned, it is important to note that some discoveries are made by **serendipity**, that is, by means of a fortunate accident or a lucky surprise. Penicillin was discovered when biologist Alexander Fleming accidentally left a petri dish of *Staphylococcus* bacteria open. An unwanted mold grew on the dish, killing the bacteria. The mold turned out to be *Penicillium*, and a new antibiotic was discovered. Even in the highly organized world of science, luck—when combined with an observant, curious mind—can lead to unexpected breakthroughs.

## Reporting Scientific Work

Whether scientific research is basic science or applied science, scientists must share their findings in order for other researchers to expand and build upon their discoveries. Collaboration with other scientists—when planning, conducting, and analyzing results—are all important for scientific research. For this reason, important aspects of a scientist's work are communicating with peers and disseminating results to peers. Scientists can share results by presenting them at a scientific meeting or conference, but this approach can reach only the select few who are present. Instead, most scientists present their results in peer-reviewed manuscripts that are published in scientific journals. **Peer-reviewed manuscripts** are scientific papers that are reviewed by a scientist's colleagues, or peers. These colleagues are qualified individuals, often experts in the same research area, who judge whether or not the scientist's work is suitable for publication. The process of peer review helps to ensure that the research described in a scientific paper or grant proposal is original, significant, logical, and thorough. Grant proposals, which are requests for research funding, are also subject to peer review. Scientists publish their work so other scientists can reproduce their experiments under similar or different conditions to expand on the findings. The experimental results must be consistent with the findings of other scientists.

A scientific paper is very different from creative writing. Although creativity is required to design experiments, there are fixed guidelines when it comes to presenting scientific results. First, scientific writing must be

brief, concise, and accurate. A scientific paper needs to be succinct but detailed enough to allow peers to reproduce the experiments.

The scientific paper consists of several specific sections—introduction, materials and methods, results, and discussion. This structure is sometimes called the “IMRaD” format. There are usually acknowledgment and reference sections as well as an **abstract** (a concise summary) at the beginning of the paper. There might be additional sections depending on the type of paper and the journal where it will be published; for example, some review papers require an outline.

The **introduction** starts with brief, but broad, background information about what is known in the field. A good introduction also gives the rationale of the work; it justifies the work carried out and also briefly mentions the end of the paper, where the hypothesis or research question driving the research will be presented. The introduction refers to the published scientific work of others and therefore requires citations following the style of the journal. Using the work or ideas of others without proper citation is considered **plagiarism**.

The **materials and methods** section includes a complete and accurate description of the substances used, and the method and techniques used by the researchers to gather data. The description should be thorough enough to allow another researcher to repeat the experiment and obtain similar results, but it does not have to be verbose. This section will also include information on how measurements were made and what types of calculations and statistical analyses were used to examine raw data. Although the materials and methods section gives an accurate description of the experiments, it does not discuss them.

Some journals require a results section followed by a discussion section, but it is more common to combine both. If the journal does not allow the combination of both sections, the **results** section simply narrates the findings without any further interpretation. The results are presented by means of tables or graphs, but no duplicate information should be presented. In the **discussion** section, the researcher will interpret the results, describe how variables may be related, and attempt to explain the observations. It is indispensable to conduct an extensive literature search to

put the results in the context of previously published scientific research. Therefore, proper citations are included in this section as well.

Finally, the **conclusion** section summarizes the importance of the experimental findings. While the scientific paper almost certainly answered one or more scientific questions that were stated, any good research should lead to more questions. Therefore, a well-done scientific paper leaves doors open for the researcher and others to continue and expand on the findings.

**Review articles** do not follow the IMRAD format because they do not present original scientific findings, or primary literature; instead, they summarize and comment on findings that were published as primary literature and typically include extensive reference sections.

## Section Summary

Biology is the science that studies living organisms and their interactions with one another and their environments. Science attempts to describe and understand the nature of the universe in whole or in part by rational means. Science has many fields; those fields related to the physical world and its phenomena are considered natural sciences.

Science can be basic or applied. The main goal of basic science is to expand knowledge without any expectation of short-term practical application of that knowledge. The primary goal of applied research, however, is to solve practical problems.

Two types of logical reasoning are used in science. Inductive reasoning uses particular results to produce general scientific principles. Deductive reasoning is a form of logical thinking that predicts results by applying general principles. The common thread throughout scientific research is the use of the scientific method, a step-based process that consists of making observations, defining a problem, posing hypotheses, testing these hypotheses, and drawing one or more conclusions. The testing uses proper controls. Scientists present their results in peer-reviewed scientific papers published in scientific journals. A scientific research paper consists of several well-defined sections: introduction, materials and methods, results,

and, finally, a concluding discussion. Review papers summarize the research done in a particular field over a period of time.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) In the example below, the scientific method is used to solve an everyday problem. Order the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items). Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation
  2. Question
  3. Hypothesis (answer)
  4. Prediction
  5. Experiment
  6. Result
- 
- a. There is something wrong with the electrical outlet.
  - b. If something is wrong with the outlet, my coffeemaker also won't work when plugged into it.
  - c. My toaster doesn't toast my bread.
  - d. I plug my coffee maker into the outlet.
  - e. My coffeemaker works.
  - f. Why doesn't my toaster work?

---

#### Solution:

[\[link\]](#) 1: C; 2: F; 3: A; 4: B; 5: D; 6: E. The original hypothesis is incorrect, as the coffeemaker works when plugged into the outlet. Alternative hypotheses include that the toaster might be broken or that the toaster wasn't turned on.

### Exercise:

**Problem:**

[\[link\]](#) Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.
2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA, separate into daughter cells during cell division. Therefore, DNA is the genetic material.
4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

---

**Solution:**

[\[link\]](#) 1: inductive; 2: deductive; 3: deductive; 4: inductive.

**Review Questions****Exercise:**

**Problem:** The first forms of life on Earth were \_\_\_\_\_.

- a. plants
- b. microorganisms
- c. birds
- d. dinosaurs

---

**Solution:**

B

**Exercise:**

**Problem:**

A suggested and testable explanation for an event is called a \_\_\_\_\_.

- a. hypothesis
- b. variable
- c. theory
- d. control

---

**Solution:**

A

**Exercise:**

**Problem:**

Which of the following sciences is not considered a natural science?

- a. biology
- b. astronomy
- c. physics
- d. computer science

---

**Solution:**

D

**Exercise:**

**Problem:**

The type of logical thinking that uses related observations to arrive at a general conclusion is called \_\_\_\_\_.

- a. deductive reasoning
  - b. the scientific method
  - c. hypothesis-based science
  - d. inductive reasoning
- 

**Solution:**

D

**Exercise:**

**Problem:**

The process of \_\_\_\_\_ helps to ensure that a scientist's research is original, significant, logical, and thorough.

- a. publication
  - b. public speaking
  - c. peer review
  - d. the scientific method
- 

**Solution:**

C

**Exercise:**

**Problem:**

A person notices that her houseplants that are regularly exposed to music seem to grow more quickly than those in rooms with no music. As a result, she determines that plants grow better when exposed to music. This example most closely resembles which type of reasoning?

- a. inductive reasoning
- b. deductive reasoning
- c. neither, because no hypothesis was made
- d. both inductive and deductive reasoning



---

**Solution:**

A

**Free Response****Exercise:****Problem:**

Although the scientific method is used by most of the sciences, it can also be applied to everyday situations. Think about a problem that you may have at home, at school, or with your car, and apply the scientific method to solve it.

---

**Solution:**

Answers will vary, but should apply the steps of the scientific method. One possibility could be a car which doesn't start. The hypothesis could be that the car doesn't start because the battery is dead. The experiment would be to change the battery or to charge the battery and then check whether the car starts or not. If it starts, the problem was due to the battery, and the hypothesis is accepted.

**Exercise:****Problem:**

Give an example of how applied science has had a direct effect on your daily life.

---

**Solution:**

Answers will vary. One example of how applied science has had a direct effect on daily life is the presence of vaccines. Vaccines to prevent diseases such as polio, measles, tetanus, and even influenza affect daily life by contributing to individual and societal health.

**Exercise:****Problem:**

Name two topics that are likely to be studied by biologists, and two areas of scientific study that would fall outside the realm of biology.

---

**Solution:**

Answers will vary. Topics that fall inside the area of biological study include how diseases affect human bodies, how pollution impacts a species' habitat, and how plants respond to their environments. Topics that fall outside of biology (the "study of life") include how metamorphic rock is formed and how planetary orbits function.

**Exercise:****Problem:**

Thinking about the topic of cancer, write a basic science question and an applied science question that a researcher interested in this topic might ask

---

**Solution:**

Answers will vary. Basic science: What evolutionary purpose might cancer serve? Applied science: What strategies might be found to prevent cancer from reproducing at the cellular level?

**Glossary**

abstract

opening section of a scientific paper that summarizes the research and conclusions

applied science

form of science that aims to solve real-world problems

basic science

science that seeks to expand knowledge and understanding regardless of the short-term application of that knowledge

biology

the study of living organisms and their interactions with one another and their environments

conclusion

section of a scientific paper that summarizes the importance of the experimental findings

control

part of an experiment that does not change during the experiment

deductive reasoning

form of logical thinking that uses a general inclusive statement to forecast specific results

descriptive science

(also, discovery science) form of science that aims to observe, explore, and investigate

discussion

section of a scientific paper in which the author interprets experimental results, describes how variables may be related, and attempts to explain the phenomenon in question

falsifiable

able to be disproven by experimental results

hypothesis

suggested explanation for an observation, which can be tested

hypothesis-based science

form of science that begins with a specific question and potential testable answers

inductive reasoning

form of logical thinking that uses related observations to arrive at a general conclusion

introduction

opening section of a scientific paper, which provides background information about what was known in the field prior to the research reported in the paper

life science

field of science, such as biology, that studies living things

materials and methods

section of a scientific paper that includes a complete description of the substances, methods, and techniques used by the researchers to gather data

natural science

field of science that is related to the physical world and its phenomena and processes

peer-reviewed manuscript

scientific paper that is reviewed by a scientist's colleagues who are experts in the field of study

physical science

field of science, such as geology, astronomy, physics, and chemistry, that studies nonliving matter

plagiarism

using other people's work or ideas without proper citation, creating the false impression that those are the author's original ideas

results

section of a scientific paper in which the author narrates the experimental findings and presents relevant figures, pictures, diagrams, graphs, and tables, without any further interpretation

review article

paper that summarizes and comments on findings that were published as primary literature

science

knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method

scientific method

method of research with defined steps that include observation, formulation of a hypothesis, testing, and confirming or falsifying the hypothesis

serendipity

fortunate accident or a lucky surprise

theory

tested and confirmed explanation for observations or phenomena

variable

part of an experiment that the experimenter can vary or change

## Themes and Concepts of Biology

By the end of this section, you will be able to:

- Identify and describe the properties of life
- Describe the levels of organization among living things
- Recognize and interpret a phylogenetic tree
- List examples of different sub disciplines in biology

Biology is the science that studies life, but what exactly is life? This may sound like a silly question with an obvious response, but it is not always easy to define life. For example, a branch of biology called virology studies viruses, which exhibit some of the characteristics of living entities but lack others. It turns out that although viruses can attack living organisms, cause diseases, and even reproduce, they do not meet the criteria that biologists use to define life. Consequently, virologists are not biologists, strictly speaking. Similarly, some biologists study the early molecular evolution that gave rise to life; since the events that preceded life are not biological events, these scientists are also excluded from biology in the strict sense of the term.

From its earliest beginnings, biology has wrestled with three questions: What are the shared properties that make something “alive”? And once we know something is alive, how do we find meaningful levels of organization in its structure? And, finally, when faced with the remarkable diversity of life, how do we organize the different kinds of organisms so that we can better understand them? As new organisms are discovered every day, biologists continue to seek answers to these and other questions.

## Properties of Life

All living organisms share several key characteristics or functions: order, sensitivity or response to the environment, reproduction, adaptation, growth and development, regulation, homeostasis, energy processing, and evolution. When viewed together, these nine characteristics serve to define life.

## Order



A toad represents a highly organized structure consisting of cells, tissues, organs, and organ systems. (credit: “Ivengo”/Wikimedia Commons)

Organisms are highly organized, coordinated structures that consist of one or more cells. Even very simple, single-celled organisms are remarkably complex: inside each cell, atoms make up molecules; these in turn make up cell organelles and other cellular inclusions. In multicellular organisms ([link](#)), similar cells form tissues. Tissues, in turn, collaborate to create organs (body structures with a distinct function). Organs work together to form organ systems.

## Sensitivity or Response to Stimuli



The leaves of this sensitive plant (*Mimosa pudica*) will instantly droop and fold when touched. After a few minutes, the plant returns to normal. (credit: Alex Lomas)

Organisms respond to diverse stimuli. For example, plants can bend toward a source of light, climb on fences and walls, or respond to touch ([\[link\]](#)). Even tiny bacteria can move toward or away from chemicals (a process called *chemotaxis*) or light (*phototaxis*). Movement toward a stimulus is considered a positive response, while movement away from a stimulus is considered a negative response.

**Note:**

Link to Learning





Watch [this video](#) to see how plants respond to a stimulus—from opening to light, to wrapping a tendril around a branch, to capturing prey.

## **Reproduction**

Single-celled organisms reproduce by first duplicating their DNA, and then dividing it equally as the cell prepares to divide to form two new cells. Multicellular organisms often produce specialized reproductive germline cells that will form new individuals. When reproduction occurs, genes containing DNA are passed along to an organism's offspring. These genes ensure that the offspring will belong to the same species and will have similar characteristics, such as size and shape.

## **Growth and Development**

Organisms grow and develop following specific instructions coded for by their genes. These genes provide instructions that will direct cellular growth and development, ensuring that a species' young ([link](#)) will grow up to exhibit many of the same characteristics as its parents.



Although no two look alike, these kittens have inherited genes from both parents and share many of the same characteristics. (credit: Rocky Mountain Feline Rescue)

## **Regulation**

Even the smallest organisms are complex and require multiple regulatory mechanisms to coordinate internal functions, respond to stimuli, and cope with environmental stresses. Two examples of internal functions regulated in an organism are nutrient transport and blood flow. Organs (groups of tissues working together) perform specific functions, such as carrying oxygen throughout the body, removing wastes, delivering nutrients to every cell, and cooling the body.

## **Homeostasis**



Polar bears (*Ursus maritimus*) and other mammals living in ice-covered regions maintain their body temperature by generating heat and reducing heat loss through thick fur and a dense layer of fat under their skin. (credit: “longhorndave”/Flickr)

In order to function properly, cells need to have appropriate conditions such as proper temperature, pH, and appropriate concentration of diverse chemicals. These conditions may, however, change from one moment to the next. Organisms are able to maintain internal conditions within a narrow range almost constantly, despite environmental changes, through **homeostasis** (literally, “steady state”)—the ability of an organism to maintain constant internal conditions. For example, an organism needs to regulate body temperature through a process known as thermoregulation. Organisms that live in cold climates, such as the polar bear ([\[link\]](#)), have body structures that help them withstand low temperatures and conserve body heat. Structures that aid in this type of insulation include fur, feathers, blubber, and fat. In hot climates, organisms have methods (such as perspiration in humans or panting in dogs) that help them to shed excess body heat.

## Energy Processing



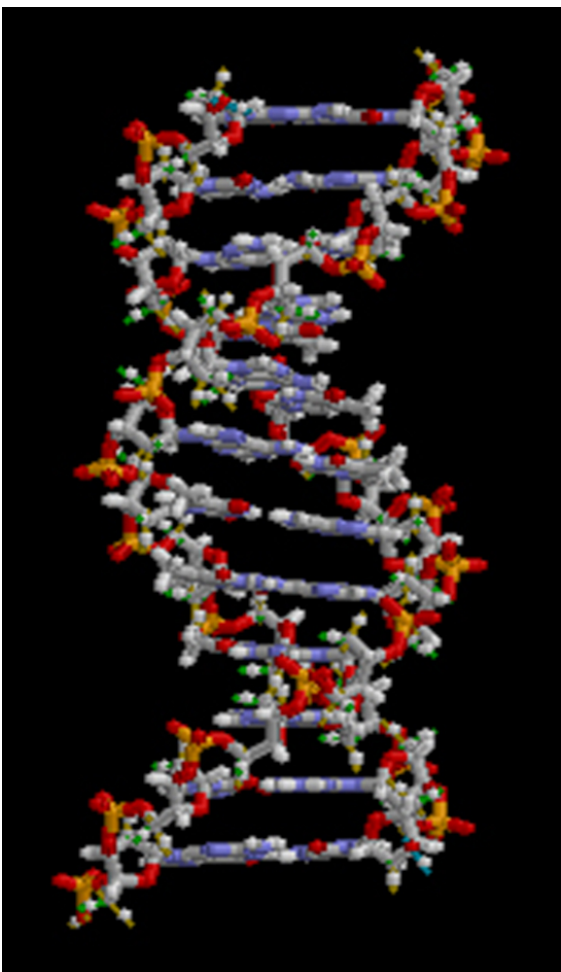
The California condor (*Gymnogyps californianus*) uses chemical energy derived from food to power flight. California condors are an endangered species; this bird has a wing tag that helps biologists identify the individual. (credit: Pacific Southwest Region U.S. Fish and Wildlife Service)

All organisms use a source of energy for their metabolic activities. Some organisms capture energy from the sun and convert it into chemical energy

in food; others use chemical energy in molecules they take in as food ([link](#)).

## Levels of Organization of Living Things

Living things are highly organized and structured, following a hierarchy that can be examined on a scale from small to large. The **atom** is the smallest and most fundamental unit of matter. It consists of a nucleus surrounded by electrons. Atoms form molecules. A **molecule** is a chemical structure consisting of at least two atoms held together by one or more chemical bonds. Many molecules that are biologically important are **macromolecules**, large molecules that are typically formed by polymerization (a polymer is a large molecule that is made by combining smaller units called monomers, which are simpler than macromolecules). An example of a macromolecule is deoxyribonucleic acid (DNA) ([link](#)), which contains the instructions for the structure and functioning of all living organisms.



All molecules, including this DNA molecule, are composed of atoms. (credit: “brian0918”/Wikimedia Commons)

**Note:**

Link to Learning



Watch [this video](#) that animates the three-dimensional structure of the DNA molecule shown in [\[link\]](#).

Some cells contain aggregates of macromolecules surrounded by membranes; these are called **organelles**. Organelles are small structures that exist within cells. Examples of organelles include mitochondria and chloroplasts, which carry out indispensable functions: mitochondria produce energy to power the cell, while chloroplasts enable green plants to utilize the energy in sunlight to make sugars. All living things are made of cells; the **cell** itself is the smallest fundamental unit of structure and function in living organisms. (This requirement is why viruses are not considered living: they are not made of cells. To make new viruses, they have to invade and hijack the reproductive mechanism of a living cell; only then can they obtain the materials they need to reproduce.) Some organisms consist of a single cell and others are multicellular. Cells are classified as prokaryotic or eukaryotic. **Prokaryotes** are single-celled or colonial organisms that do not have membrane-bound nuclei; in contrast, the cells of **eukaryotes** do have membrane-bound organelles and a membrane-bound nucleus.

In larger organisms, cells combine to make **tissues**, which are groups of similar cells carrying out similar or related functions. **Organs** are collections of tissues grouped together performing a common function. Organs are present not only in animals but also in plants. An **organ system** is a higher level of organization that consists of functionally related organs. Mammals have many organ systems. For instance, the circulatory system transports blood through the body and to and from the lungs; it includes organs such as the heart and blood vessels. **Organisms** are individual living entities. For example, each tree in a forest is an organism. Single-celled

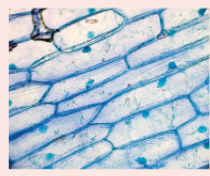
prokaryotes and single-celled eukaryotes are also considered organisms and are typically referred to as microorganisms.

All the individuals of a species living within a specific area are collectively called a **population**. For example, a forest may include many pine trees. All of these pine trees represent the population of pine trees in this forest. Different populations may live in the same specific area. For example, the forest with the pine trees includes populations of flowering plants and also insects and microbial populations. A **community** is the sum of populations inhabiting a particular area. For instance, all of the trees, flowers, insects, and other populations in a forest form the forest's community. The forest itself is an ecosystem. An **ecosystem** consists of all the living things in a particular area together with the abiotic, non-living parts of that environment such as nitrogen in the soil or rain water. At the highest level of organization ([\[link\]](#)), the **biosphere** is the collection of all ecosystems, and it represents the zones of life on earth. It includes land, water, and even the atmosphere to a certain extent.

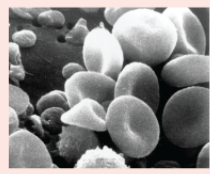
**Note:**

Art Connection

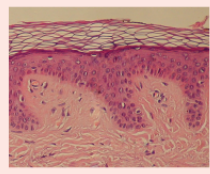




**Organelles:** The nucleus, dyed blue in these onion cells, is an example of an organelle.



**Cells:** Human blood cells.



**Tissues:** Human skin tissue.



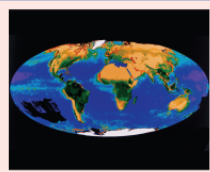
**Organs and Organ Systems:** Organs, such as the stomach and intestine, make up the human digestive system.



**Organisms, Populations, and Communities:** In a forest, each pine tree is an organism. Together, all the pine trees make up a population. All the plant and animal species in the forest comprise a community.



**Ecosystems:** This coastal ecosystem in the southeastern United States includes living organisms and the environment in which they live.



**The Biosphere:** Encompasses all the ecosystems on Earth.

The biological levels of organization of living things are shown. From a single organelle to the

entire biosphere, living organisms are parts of a highly structured hierarchy. (credit “organelles”: modification of work by Umberto Salvagnin; credit “cells”: modification of work by Bruce Wetzel, Harry Schaefer/ National Cancer Institute; credit “tissues”: modification of work by Kilbad; Fama Clamosa; Mikael Häggström; credit “organs”: modification of work by Mariana Ruiz Villareal; credit “organisms”: modification of work by "Crystal"/Flickr; credit “ecosystems”: modification of work by US Fish and Wildlife Service Headquarters; credit “biosphere”: modification of work by NASA)

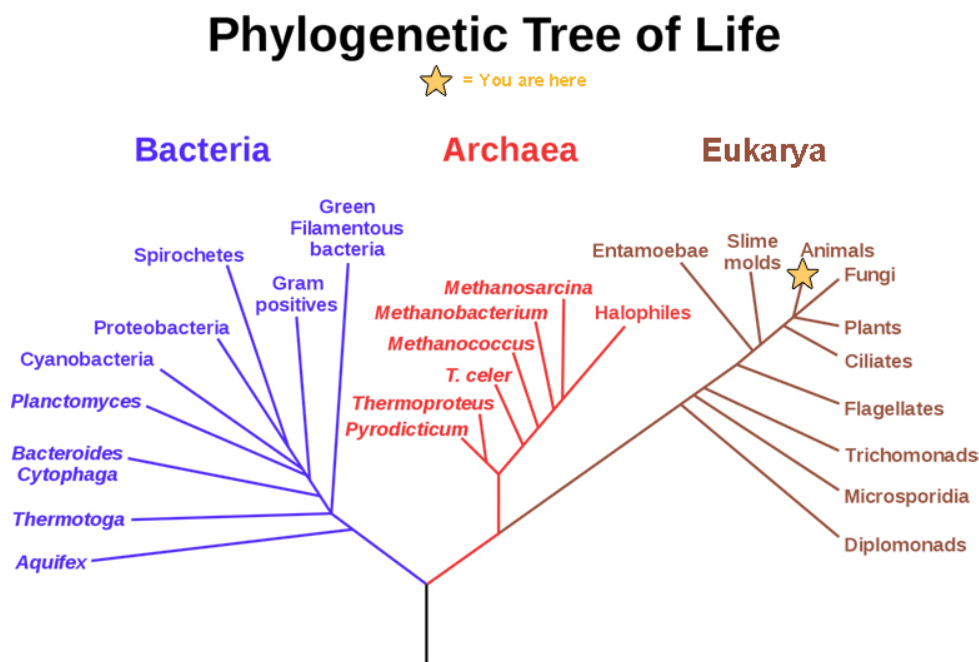
Which of the following statements is false?

- a. Tissues exist within organs which exist within organ systems.
- b. Communities exist within populations which exist within ecosystems.
- c. Organelles exist within cells which exist within tissues.
- d. Communities exist within ecosystems which exist in the biosphere.

## The Diversity of Life

The fact that biology, as a science, has such a broad scope has to do with the tremendous diversity of life on earth. The source of this diversity is **evolution**, the process of gradual change during which new species arise from older species. Evolutionary biologists study the evolution of living things in everything from the microscopic world to ecosystems.

The evolution of various life forms on Earth can be summarized in a phylogenetic tree ([link](#)). A **phylogenetic tree** is a diagram showing the evolutionary relationships among biological species based on similarities and differences in genetic or physical traits or both. A phylogenetic tree is composed of nodes and branches. The internal nodes represent ancestors and are points in evolution when, based on scientific evidence, an ancestor is thought to have diverged to form two new species. The length of each branch is proportional to the time elapsed since the split.



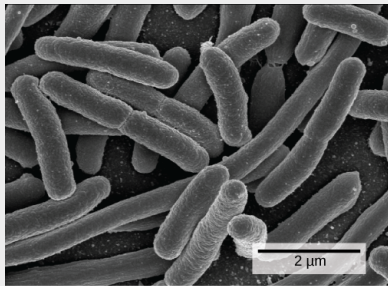
This phylogenetic tree was constructed by microbiologist Carl Woese using data obtained from sequencing ribosomal RNA genes. The tree shows the separation of living organisms into three domains: Bacteria, Archaea, and Eukarya. Bacteria and Archaea are prokaryotes, single-celled organisms lacking intracellular organelles. (credit: Eric Gaba; NASA Astrobiology Institute)

**Note:****Evolution Connection****Carl Woese and the Phylogenetic Tree**

In the past, biologists grouped living organisms into five kingdoms: animals, plants, fungi, protists, and bacteria. The organizational scheme was based mainly on physical features, as opposed to physiology, biochemistry, or molecular biology, all of which are used by modern systematics. The pioneering work of American microbiologist Carl Woese in the early 1970s has shown, however, that life on Earth has evolved along three lineages, now called domains—Bacteria, Archaea, and Eukarya. The first two are prokaryotic cells with microbes that lack membrane-enclosed nuclei and organelles. The third domain contains the eukaryotes and includes unicellular microorganisms together with the four original kingdoms (excluding bacteria). Woese defined Archaea as a new domain, and this resulted in a new taxonomic tree ([\[link\]](#)). Many organisms belonging to the Archaea domain live under extreme conditions and are called extremophiles. To construct his tree, Woese used genetic relationships rather than similarities based on morphology (shape). Woese's tree was constructed from comparative sequencing of the genes that are universally distributed, present in every organism, and conserved (meaning that these genes have remained essentially unchanged throughout evolution). Woese's approach was revolutionary because comparisons of physical features are insufficient to differentiate between the prokaryotes that appear fairly similar in spite of their tremendous biochemical diversity and genetic variability ([\[link\]](#)). The comparison of homologous DNA and

RNA sequences provided Woese with a sensitive device that revealed the extensive variability of prokaryotes, and which justified the separation of the prokaryotes into two domains: bacteria and archaea.

These images represent different domains. The (a) bacteria in this micrograph belong to Domain Bacteria, while the (b) extremophiles (not visible) living in this hot vent belong to Domain Archaea. Both the (c) sunflower and (d) lion are part of Domain Eukarya. (credit a: modification of work by Drew March; credit b: modification of work by Steve Jurvetson; credit c: modification of work by Michael Arrighi; credit d: modification of work by Leszek Leszcynski)



(a)



(b)



(c)



(d)

## Branches of Biological Study

The scope of biology is broad and therefore contains many branches and subdisciplines. Biologists may pursue one of those subdisciplines and work

in a more focused field. For instance, **molecular biology** and **biochemistry** study biological processes at the molecular and chemical level, including interactions among molecules such as DNA, RNA, and proteins, as well as the way they are regulated. **Microbiology**, the study of microorganisms, is the study of the structure and function of single-celled organisms. It is quite a broad branch itself, and depending on the subject of study, there are also microbial physiologists, ecologists, and geneticists, among others.

**Note:****Career Connection****Forensic Scientist**

Forensic science is the application of science to answer questions related to the law. Biologists as well as chemists and biochemists can be forensic scientists. Forensic scientists provide scientific evidence for use in courts, and their job involves examining trace materials associated with crimes. Interest in forensic science has increased in the last few years, possibly because of popular television shows that feature forensic scientists on the job. Also, the development of molecular techniques and the establishment of DNA databases have expanded the types of work that forensic scientists can do. Their job activities are primarily related to crimes against people such as murder, rape, and assault. Their work involves analyzing samples such as hair, blood, and other body fluids and also processing DNA ([\[link\]](#)) found in many different environments and materials. Forensic scientists also analyze other biological evidence left at crime scenes, such as insect larvae or pollen grains. Students who want to pursue careers in forensic science will most likely be required to take chemistry and biology courses as well as some intensive math courses.



This forensic scientist works in a DNA extraction room at the U.S. Army Criminal Investigation Laboratory at Fort Gillem, GA. (credit: United States Army CID Command Public Affairs)

Another field of biological study, **neurobiology**, studies the biology of the nervous system, and although it is considered a branch of biology, it is also recognized as an interdisciplinary field of study known as neuroscience. Because of its interdisciplinary nature, this subdiscipline studies different functions of the nervous system using molecular, cellular, developmental, medical, and computational approaches.





Researchers work on excavating dinosaur fossils at a site in Castellón, Spain. (credit: Mario Modesto)

**Paleontology**, another branch of biology, uses fossils to study life's history ([link](#)). **Zoology** and **botany** are the study of animals and plants, respectively. Biologists can also specialize as biotechnologists, ecologists, or physiologists, to name just a few areas. This is just a small sample of the many fields that biologists can pursue.

Biology is the culmination of the achievements of the natural sciences from their inception to today. Excitingly, it is the cradle of emerging sciences, such as the biology of brain activity, genetic engineering of custom organisms, and the biology of evolution that uses the laboratory tools of molecular biology to retrace the earliest stages of life on earth. A scan of news headlines—whether reporting on immunizations, a newly discovered species, sports doping, or a genetically-modified food—demonstrates the way biology is active in and important to our everyday world.

## Section Summary



Biology is the science of life. All living organisms share several key properties such as order, sensitivity or response to stimuli, reproduction, growth and development, regulation, homeostasis, and energy processing. Living things are highly organized parts of a hierarchy that includes atoms, molecules, organelles, cells, tissues, organs, and organ systems. Organisms, in turn, are grouped as populations, communities, ecosystems, and the biosphere. The great diversity of life today evolved from less-diverse ancestral organisms over billions of years. A diagram called a phylogenetic tree can be used to show evolutionary relationships among organisms.

Biology is very broad and includes many branches and subdisciplines. Examples include molecular biology, microbiology, neurobiology, zoology, and botany, among others.

## Art Connections

### Exercise:

**Problem:** [\[link\]](#) Which of the following statements is false?

- a. Tissues exist within organs which exist within organ systems.
- b. Communities exist within populations which exist within ecosystems.
- c. Organelles exist within cells which exist within tissues.
- d. Communities exist within ecosystems which exist in the biosphere.

---

### Solution:

[\[link\]](#) Communities exist within populations which exist within ecosystems.

## Review Questions

### Exercise:

**Problem:**

The smallest unit of biological structure that meets the functional requirements of “living” is the \_\_\_\_\_.

- a. organ
- b. organelle
- c. cell
- d. macromolecule

---

**Solution:**

C

**Exercise:**

**Problem:** Viruses are not considered living because they \_\_\_\_\_.

- a. are not made of cells
- b. lack cell nuclei
- c. do not contain DNA or RNA
- d. cannot reproduce

---

**Solution:**

A

**Exercise:**

**Problem:**

The presence of a membrane-enclosed nucleus is a characteristic of \_\_\_\_\_.

- a. prokaryotic cells
- b. eukaryotic cells
- c. living organisms

d. bacteria

---

**Solution:**

B

**Exercise:**

**Problem:**

A group of individuals of the same species living in the same area is called a(n) \_\_\_\_\_.

- a. family
- b. community
- c. population
- d. ecosystem

---

**Solution:**

C

**Exercise:**

**Problem:**

Which of the following sequences represents the hierarchy of biological organization from the most inclusive to the least complex level?

- a. organelle, tissue, biosphere, ecosystem, population
- b. organ, organism, tissue, organelle, molecule
- c. organism, community, biosphere, molecule, tissue, organ
- d. biosphere, ecosystem, community, population, organism

---

**Solution:**

D

**Exercise:****Problem:**

Where in a phylogenetic tree would you expect to find the organism that had evolved most recently?

- a. at the base
- b. within the branches
- c. at the nodes
- d. at the branch tips

---

**Solution:**

D

**Free Response****Exercise:****Problem:**

Select two items that biologists agree are necessary in order to consider an organism “alive.” For each, give an example of a non-living object that otherwise fits the definition of “alive,”

---

**Solution:**

Answers will vary. Layers of sedimentary rock have order but are not alive. Technology is capable of regulation but is not, of itself, alive.

**Exercise:**

**Problem:**

Consider the levels of organization of the biological world, and place each of these items in order from smallest level of organization to most encompassing: skin cell, elephant, water molecule, planet Earth, tropical rainforest, hydrogen atom, wolf pack, liver.

---

**Solution:**

Smallest level of organization to largest: hydrogen atom, water molecule, skin cell, liver, elephant, wolf pack, tropical rainforest, planet Earth

**Exercise:****Problem:**

You go for a long walk on a hot day. Give an example of a way in which homeostasis keeps your body healthy.

---

**Solution:**

During your walk, you may begin to perspire, which cools your body and helps your body to maintain a constant internal temperature. You might also become thirsty and pause long enough for a cool drink, which will help to restore the water lost during perspiration.

**Exercise:****Problem:**

Using examples, explain how biology can be studied from a microscopic approach to a global approach.

---

**Solution:**

Researchers can approach biology from the smallest to the largest, and everything in between. For instance, an ecologist may study a population of individuals, the population's community, the community's ecosystem, and the ecosystem's part in the biosphere.

When studying an individual organism, a biologist could examine the cell and its organelles, the tissues that the cells make up, the organs and their respective organ systems, and the sum total—the organism itself.

## **Glossary**

atom

smallest and most fundamental unit of matter

biochemistry

study of the chemistry of biological organisms

biosphere

collection of all the ecosystems on Earth

botany

study of plants

cell

smallest fundamental unit of structure and function in living things

community

set of populations inhabiting a particular area

ecosystem

all the living things in a particular area together with the abiotic, nonliving parts of that environment

eukaryote

organism with cells that have nuclei and membrane-bound organelles

evolution

process of gradual change during which new species arise from older species and some species become extinct

homeostasis

ability of an organism to maintain constant internal conditions

macromolecule

large molecule, typically formed by the joining of smaller molecules

microbiology

study of the structure and function of microorganisms

molecule

chemical structure consisting of at least two atoms held together by one or more chemical bonds

molecular biology

study of biological processes and their regulation at the molecular level, including interactions among molecules such as DNA, RNA, and proteins

neurobiology

study of the biology of the nervous system

organ

collection of related tissues grouped together performing a common function

organ system

level of organization that consists of functionally related interacting organs

organelle

small structures that exist within cells and carry out cellular functions

organism

individual living entity

paleontology

study of life's history by means of fossils

phylogenetic tree

diagram showing the evolutionary relationships among various biological species based on similarities and differences in genetic or physical traits or both; in essence, a hypothesis concerning evolutionary connections

population

all of the individuals of a species living within a specific area

prokaryote

single-celled organism that lacks organelles and does not have nuclei surrounded by a nuclear membrane

tissue

group of similar cells carrying out related functions

zoology

study of animals



## Introduction

class="introduction"

Foods such as bread, fruit, and cheese are rich sources of biological macromolecules . (credit: modification of work by Bengt Nyman)



Food provides the body with the nutrients it needs to survive. Many of these critical nutrients are biological macromolecules, or large molecules, necessary for life. These macromolecules (polymers) are built from

different combinations of smaller organic molecules (monomers). What specific types of biological macromolecules do living things require? How are these molecules formed? What functions do they serve? In this chapter, these questions will be explored.

## Synthesis of Biological Macromolecules

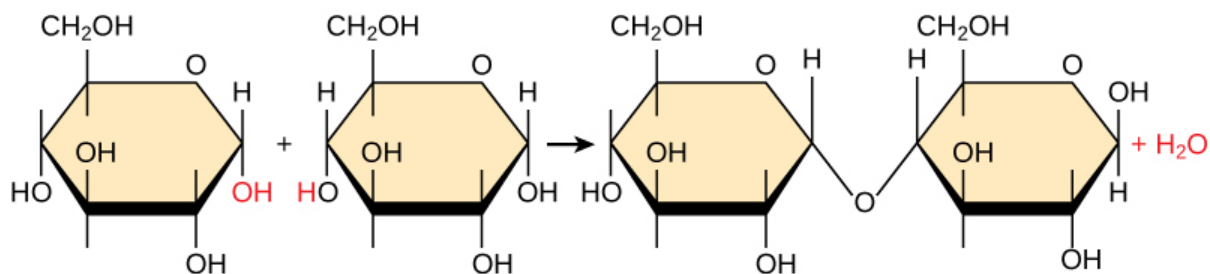
By the end of this section, you will be able to:

- Understand the synthesis of macromolecules
- Explain dehydration (or condensation) and hydrolysis reactions

As you've learned, **biological macromolecules** are large molecules, necessary for life, that are built from smaller organic molecules. There are four major classes of biological macromolecules (carbohydrates, lipids, proteins, and nucleic acids); each is an important cell component and performs a wide array of functions. Combined, these molecules make up the majority of a cell's dry mass (recall that water makes up the majority of its complete mass). Biological macromolecules are organic, meaning they contain carbon. In addition, they may contain hydrogen, oxygen, nitrogen, and additional minor elements.

### Dehydration Synthesis

Most macromolecules are made from single subunits, or building blocks, called **monomers**. The monomers combine with each other using covalent bonds to form larger molecules known as **polymers**. In doing so, monomers release water molecules as byproducts. This type of reaction is known as **dehydration synthesis**, which means “to put together while losing water.”

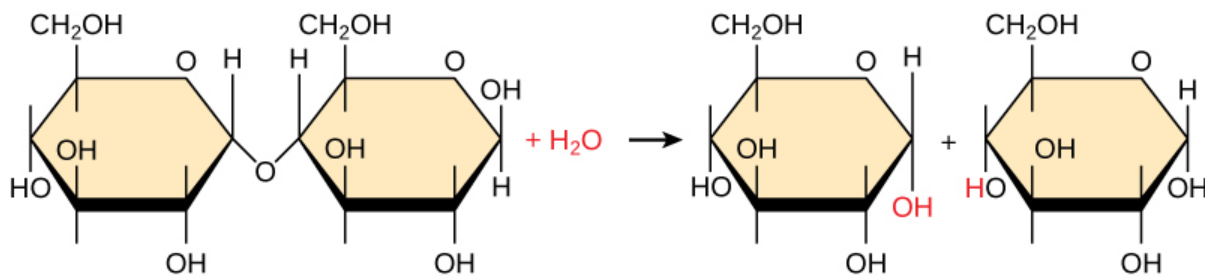


In the dehydration synthesis reaction depicted above, two molecules of glucose are linked together to form the disaccharide maltose. In the process, a water molecule is formed.

In a dehydration synthesis reaction ([\[link\]](#)), the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a molecule of water. At the same time, the monomers share electrons and form covalent bonds. As additional monomers join, this chain of repeating monomers forms a polymer. Different types of monomers can combine in many configurations, giving rise to a diverse group of macromolecules. Even one kind of monomer can combine in a variety of ways to form several different polymers: for example, glucose monomers are the constituents of starch, glycogen, and cellulose.

## Hydrolysis

Polymers are broken down into monomers in a process known as hydrolysis, which means “to split water,” a reaction in which a water molecule is used during the breakdown ([\[link\]](#)). During these reactions, the polymer is broken into two components: one part gains a hydrogen atom ( $H^+$ ) and the other gains a hydroxyl molecule ( $OH^-$ ) from a split water molecule.



In the hydrolysis reaction shown here, the disaccharide maltose is broken down to form two glucose monomers with the addition of a water molecule. Note that this reaction is the reverse of the synthesis reaction shown in [\[link\]](#).

Dehydration and **hydrolysis reactions** are catalyzed, or “sped up,” by specific enzymes; dehydration reactions involve the formation of new bonds, requiring energy, while hydrolysis reactions break bonds and release energy. These reactions are similar for most macromolecules, but each monomer and polymer reaction is specific for its class. For example, in our bodies, food is hydrolyzed, or broken down, into smaller molecules by catalytic enzymes in the digestive system. This allows for easy absorption of nutrients by cells in the intestine. Each macromolecule is broken down by a specific enzyme. For instance, carbohydrates are broken down by amylase, sucrase, lactase, or maltase. Proteins are broken down by the enzymes pepsin and peptidase, and by hydrochloric acid. Lipids are broken down by lipases. Breakdown of these macromolecules provides energy for cellular activities.

**Note:**

Link to Learning



Visit [this site](#) to see visual representations of dehydration synthesis and hydrolysis.

## Section Summary

Proteins, carbohydrates, nucleic acids, and lipids are the four major classes of biological macromolecules—large molecules necessary for life that are built from smaller organic molecules. Macromolecules are made up of single units known as monomers that are joined by covalent bonds to form larger polymers. The polymer is more than the sum of its parts: it acquires

new characteristics, and leads to an osmotic pressure that is much lower than that formed by its ingredients; this is an important advantage in the maintenance of cellular osmotic conditions. A monomer joins with another monomer with the release of a water molecule, leading to the formation of a covalent bond. These types of reactions are known as dehydration or condensation reactions. When polymers are broken down into smaller units (monomers), a molecule of water is used for each bond broken by these reactions; such reactions are known as hydrolysis reactions. Dehydration and hydrolysis reactions are similar for all macromolecules, but each monomer and polymer reaction is specific to its class. Dehydration reactions typically require an investment of energy for new bond formation, while hydrolysis reactions typically release energy by breaking bonds.

## Review Questions

### Exercise:

**Problem:** Dehydration synthesis leads to formation of

- a. monomers
- b. polymers
- c. water and polymers
- d. none of the above

---

### Solution:

C

### Exercise:

**Problem:**

During the breakdown of polymers, which of the following reactions takes place?

- a. hydrolysis
- b. dehydration

- c. condensation
  - d. covalent bond
- 

**Solution:**

A

## Free Response

**Exercise:**

**Problem:** Why are biological macromolecules considered organic?

---

**Solution:**

Biological macromolecules are organic because they contain carbon.

**Exercise:**

**Problem:**

What role do electrons play in dehydration synthesis and hydrolysis?

---

**Solution:**

In a dehydration synthesis reaction, the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a molecule of water. This creates an opening in the outer shells of atoms in the monomers, which can share electrons and form covalent bonds.

## Glossary

biological macromolecule

large molecule necessary for life that is built from smaller organic molecules

dehydration synthesis

(also, condensation) reaction that links monomer molecules together, releasing a molecule of water for each bond formed

hydrolysis

reaction causes breakdown of larger molecules into smaller molecules with the utilization of water

monomer

smallest unit of larger molecules called polymers

polymer

chain of monomer residues that is linked by covalent bonds;

polymerization is the process of polymer formation from monomers by condensation



## Carbohydrates

By the end of this section, you will be able to:

- Discuss the role of carbohydrates in cells and in the extracellular materials of animals and plants
- Explain the classifications of carbohydrates
- List common monosaccharides, disaccharides, and polysaccharides

Most people are familiar with carbohydrates, one type of macromolecule, especially when it comes to what we eat. To lose weight, some individuals adhere to “low-carb” diets. Athletes, in contrast, often “carb-load” before important competitions to ensure that they have enough energy to compete at a high level. Carbohydrates are, in fact, an essential part of our diet; grains, fruits, and vegetables are all natural sources of carbohydrates. Carbohydrates provide energy to the body, particularly through glucose, a simple sugar that is a component of **starch** and an ingredient in many staple foods. Carbohydrates also have other important functions in humans, animals, and plants.

## Molecular Structures

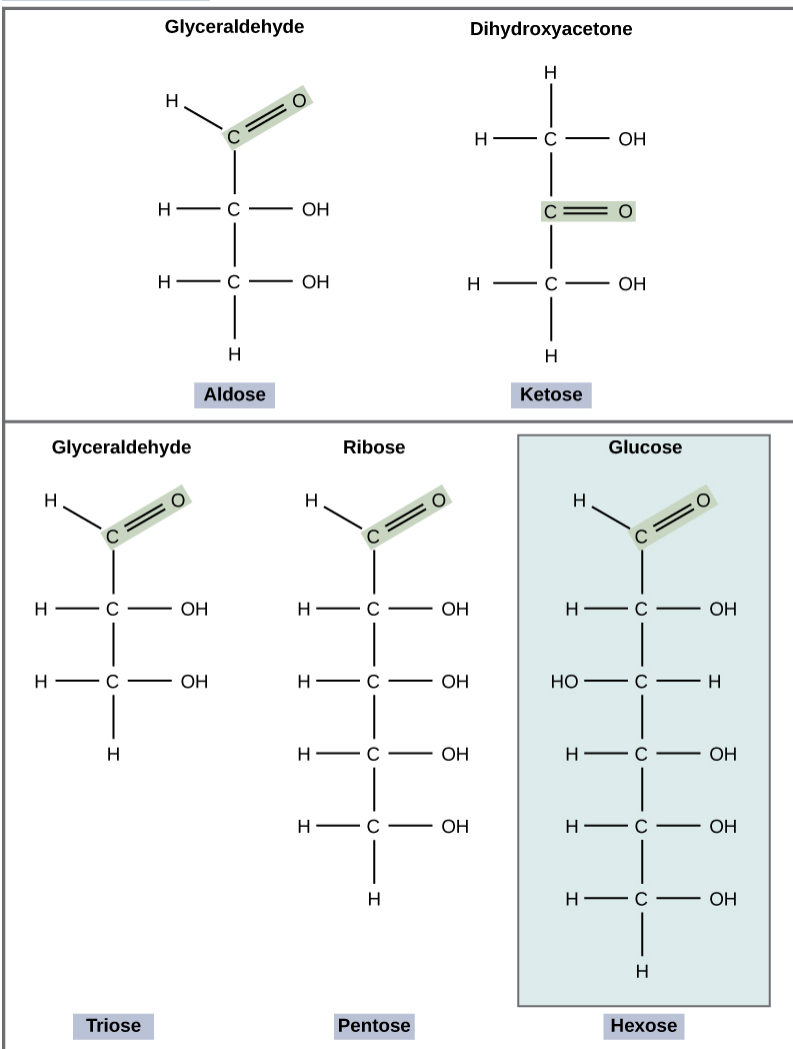
**Carbohydrates** can be represented by the stoichiometric formula  $(\text{CH}_2\text{O})_n$ , where  $n$  is the number of carbons in the molecule. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. This formula also explains the origin of the term “carbohydrate”: the components are carbon (“carbo”) and the components of water (hence, “hydrate”). Carbohydrates are classified into three subtypes: monosaccharides, disaccharides, and polysaccharides.

## Monosaccharides

**Monosaccharides** (mono- = “one”; sacchar- = “sweet”) are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbons usually ranges from three to seven. Most monosaccharide names end with the suffix -ose. If the sugar has an aldehyde group (the functional group with the structure  $\text{R-CHO}$ ), it is known as an aldose, and if it has a

ketone group (the functional group with the structure  $\text{RC}(=\text{O})\text{R}'$ ), it is known as a ketose. Depending on the number of carbons in the sugar, they also may be known as trioses (three carbons), pentoses (five carbons), and or hexoses (six carbons). See [\[link\]](#) for an illustration of the monosaccharides.

#### MONOSACCHARIDES



Monosaccharides are classified based on the position of their carbonyl group and the number of carbons in the backbone. Aldoses have a carbonyl group (indicated in green) at the end of the carbon chain, and ketoses

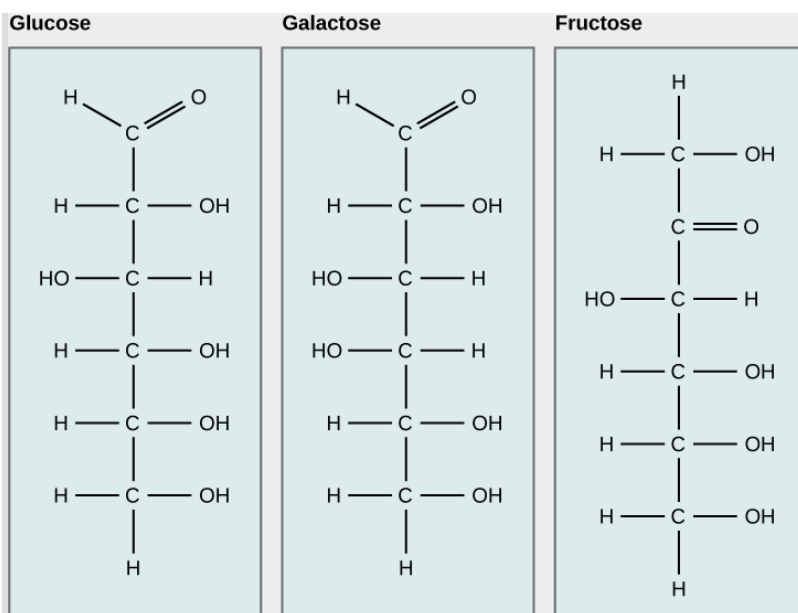
have a carbonyl group in the middle of the carbon chain. Trioses, pentoses, and hexoses have three, five, and six carbon backbones, respectively.

The chemical formula for glucose is  $C_6H_{12}O_6$ . In humans, glucose is an important source of energy. During cellular respiration, energy is released from glucose, and that energy is used to help make adenosine triphosphate (ATP). Plants synthesize glucose using carbon dioxide and water, and glucose in turn is used for energy requirements for the plant. Excess glucose is often stored as starch that is catabolized (the breakdown of larger molecules by cells) by humans and other animals that feed on plants.

Galactose (part of lactose, or milk sugar) and fructose (found in sucrose, in fruit) are other common monosaccharides. Although glucose, galactose, and fructose all have the same chemical formula ( $C_6H_{12}O_6$ ), they differ structurally and chemically (and are known as isomers) because of the different arrangement of functional groups around the asymmetric carbon; all of these monosaccharides have more than one asymmetric carbon ([link](#)).

**Note:**

Art Connection

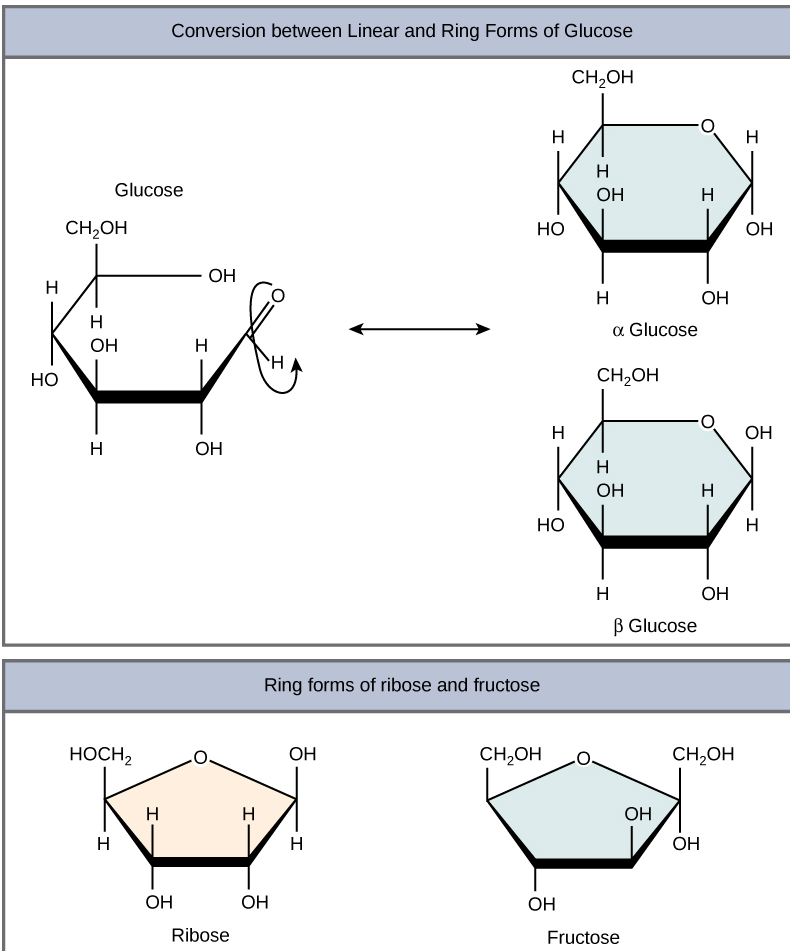


Glucose, galactose, and fructose are all hexoses. They are structural isomers, meaning they have the same chemical formula ( $\text{C}_6\text{H}_{12}\text{O}_6$ ) but a different arrangement of atoms.

What kind of sugars are these, aldose or ketose?

Glucose, galactose, and fructose are isomeric monosaccharides (hexoses), meaning they have the same chemical formula but have slightly different structures. Glucose and galactose are aldoses, and fructose is a ketose.

Monosaccharides can exist as a linear chain or as ring-shaped molecules; in aqueous solutions they are usually found in ring forms ([link](#)). Glucose in a ring form can have two different arrangements of the hydroxyl group (OH) around the anomeric carbon (carbon 1 that becomes asymmetric in the process of ring formation). If the hydroxyl group is below carbon number 1 in the sugar, it is said to be in the alpha ( $\alpha$ ) position, and if it is above the plane, it is said to be in the beta ( $\beta$ ) position.

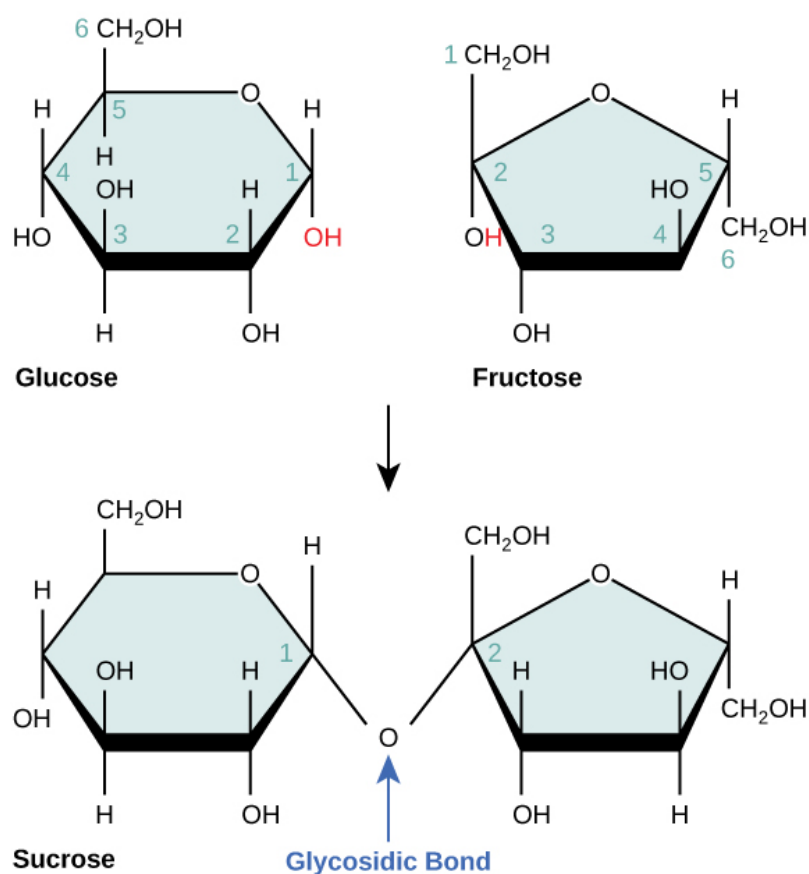


Five and six carbon monosaccharides exist in equilibrium between linear and ring forms. When the ring forms, the side chain it closes on is locked into an  $\alpha$  or  $\beta$  position.

Fructose and ribose also form rings, although they form five-membered rings as opposed to the six-membered ring of glucose.

## Disaccharides

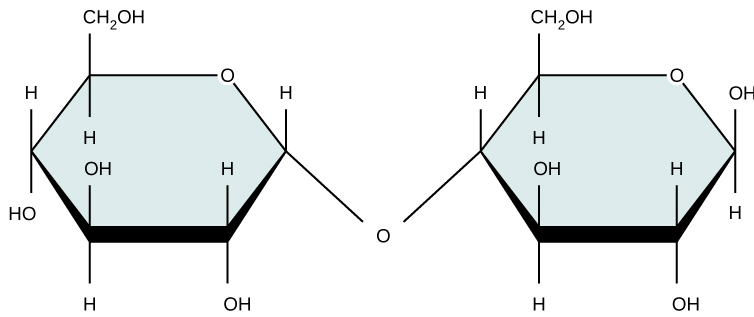
**Disaccharides** (di- = “two”) form when two monosaccharides undergo a dehydration reaction (also known as a condensation reaction or dehydration synthesis). During this process, the hydroxyl group of one monosaccharide combines with the hydrogen of another monosaccharide, releasing a molecule of water and forming a covalent bond. A covalent bond formed between a carbohydrate molecule and another molecule (in this case, between two monosaccharides) is known as a **glycosidic bond** ([\[link\]](#)). Glycosidic bonds (also called glycosidic linkages) can be of the alpha or the beta type.



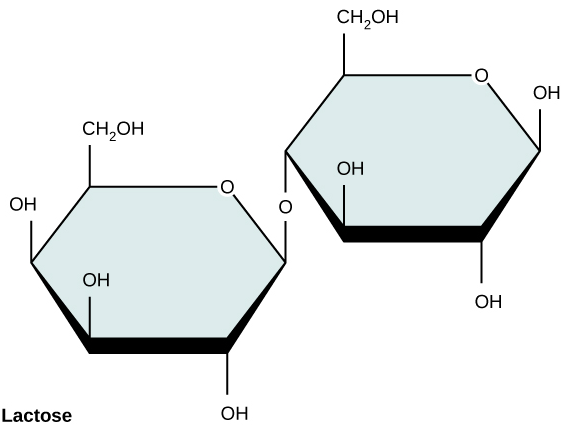
Sucrose is formed when a monomer of glucose and a monomer of fructose are joined in a dehydration reaction to form a glycosidic bond. In the process, a water molecule is lost. By convention, the carbon

atoms in a monosaccharide are numbered from the terminal carbon closest to the carbonyl group. In sucrose, a glycosidic linkage is formed between carbon 1 in glucose and carbon 2 in fructose.

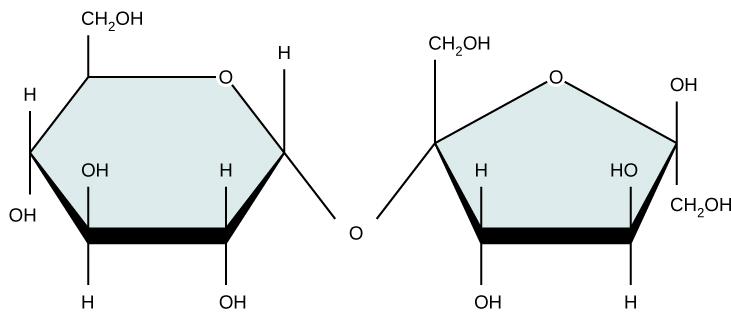
Common disaccharides include lactose, maltose, and sucrose ([\[link\]](#)). Lactose is a disaccharide consisting of the monomers glucose and galactose. It is found naturally in milk. Maltose, or malt sugar, is a disaccharide formed by a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is composed of the monomers glucose and fructose.



**Maltose**



**Lactose**



**Sucrose**

Common disaccharides include maltose (grain sugar), lactose (milk sugar), and sucrose (table sugar).

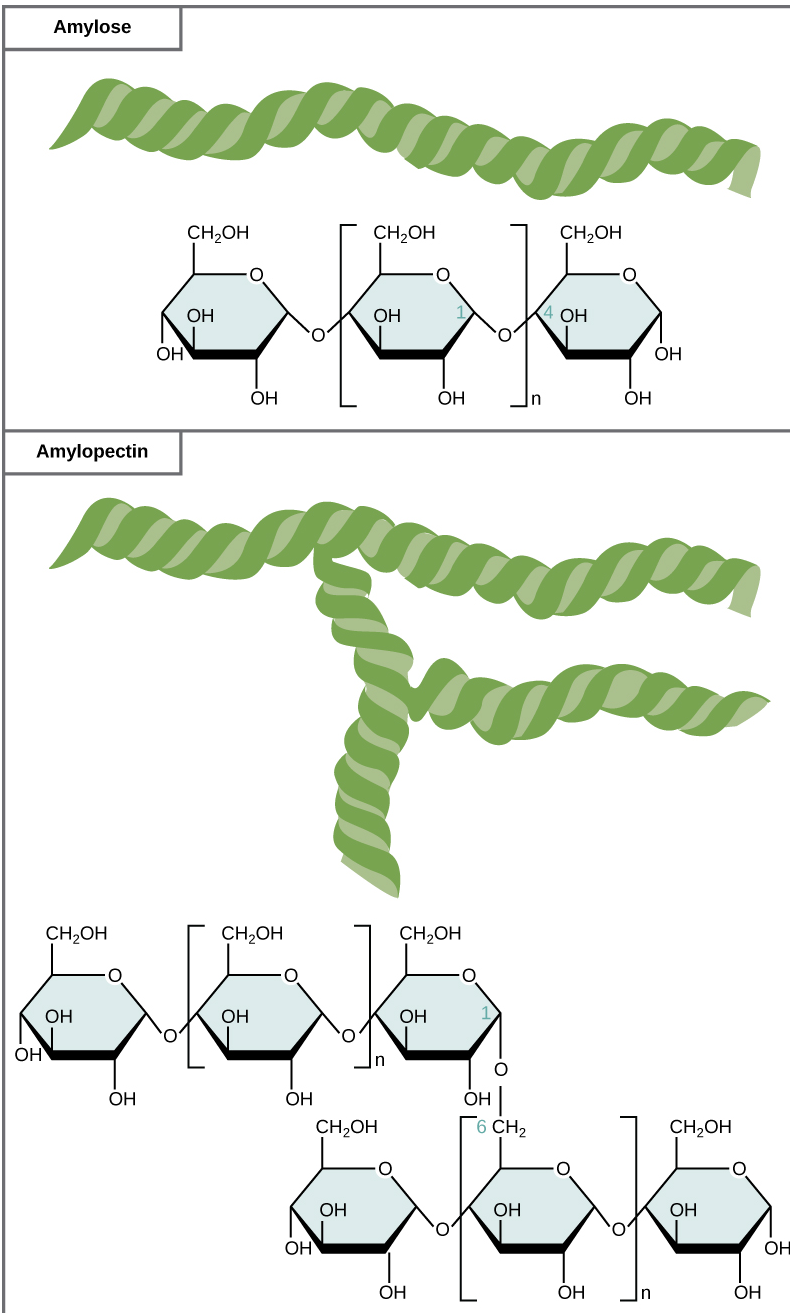
## Polysaccharides



A long chain of monosaccharides linked by glycosidic bonds is known as a **polysaccharide** (poly- = “many”). The chain may be branched or unbranched, and it may contain different types of monosaccharides. The molecular weight may be 100,000 daltons or more depending on the number of monomers joined. Starch, glycogen, cellulose, and chitin are primary examples of polysaccharides.

Starch is the stored form of sugars in plants and is made up of a mixture of amylose and amylopectin (both polymers of glucose). Plants are able to synthesize glucose, and the excess glucose, beyond the plant’s immediate energy needs, is stored as starch in different plant parts, including roots and seeds. The starch in the seeds provides food for the embryo as it germinates and can also act as a source of food for humans and animals. The starch that is consumed by humans is broken down by enzymes, such as salivary amylases, into smaller molecules, such as maltose and glucose. The cells can then absorb the glucose.

Starch is made up of glucose monomers that are joined by  $\alpha$  1-4 or  $\alpha$  1-6 glycosidic bonds. The numbers 1-4 and 1-6 refer to the carbon number of the two residues that have joined to form the bond. As illustrated in [\[link\]](#), amylose is starch formed by unbranched chains of glucose monomers (only  $\alpha$  1-4 linkages), whereas amylopectin is a branched polysaccharide ( $\alpha$  1-6 linkages at the branch points).



Amylose and amylopectin are two different forms of starch. Amylose is composed of unbranched chains of glucose monomers connected by  $\alpha$  1,4 glycosidic linkages.

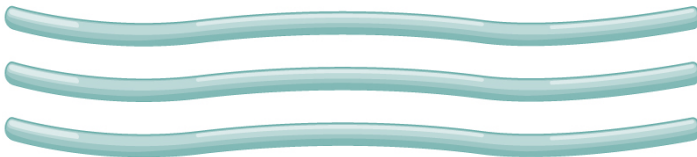
Amylopectin is composed of branched chains of glucose monomers connected by  $\alpha$  1,4 and  $\alpha$  1,6 glycosidic linkages. Because

of the way the subunits are joined, the glucose chains have a helical structure. Glycogen (not shown) is similar in structure to amylopectin but more highly branched.

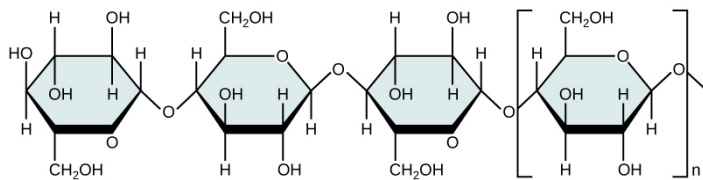
**Glycogen** is the storage form of glucose in humans and other vertebrates and is made up of monomers of glucose. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever blood glucose levels decrease, glycogen is broken down to release glucose in a process known as glycogenolysis.

**Cellulose** is the most abundant natural biopolymer. The cell wall of plants is mostly made of cellulose; this provides structural support to the cell. Wood and paper are mostly cellulosic in nature. Cellulose is made up of glucose monomers that are linked by  $\beta$  1-4 glycosidic bonds ([\[link\]](#)).

Cellulose fibers



Cellulose structure



In cellulose, glucose monomers are linked in unbranched chains by  $\beta$  1-4 glycosidic linkages. Because of the way the glucose subunits are joined, every glucose monomer is flipped relative to the next one resulting in a linear, fibrous structure.

As shown in [\[link\]](#), every other glucose monomer in cellulose is flipped over, and the monomers are packed tightly as extended long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. While the  $\beta$  1-4 linkage cannot be broken down by human digestive enzymes, herbivores such as cows, koalas, and buffalos are able, with the help of the specialized flora in their stomach, to digest plant material that is rich in cellulose and use it as a food source. In these animals, certain species of bacteria and protists reside in the rumen (part of the digestive system of herbivores) and secrete the enzyme cellulase. The appendix of grazing animals also contains bacteria that digest cellulose, giving it an important role in the digestive systems of ruminants. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal. Termites are also able to break down cellulose because of the presence of other organisms in their bodies that secrete cellulases.

Carbohydrates serve various functions in different animals. Arthropods (insects, crustaceans, and others) have an outer skeleton, called the exoskeleton, which protects their internal body parts (as seen in the bee in [\[link\]](#)). This exoskeleton is made of the biological macromolecule **chitin**, which is a polysaccharide-containing nitrogen. It is made of repeating units of N-acetyl- $\beta$ -d-glucosamine, a modified sugar. Chitin is also a major component of fungal cell walls; fungi are neither animals nor plants and form a kingdom of their own in the domain Eukarya.



Insects have a hard outer exoskeleton made of chitin, a type of polysaccharide. (credit: Louise Docker)

**Note:**

**Career Connections**

**Registered Dietitian**

Obesity is a worldwide health concern, and many diseases such as diabetes and heart disease are becoming more prevalent because of obesity. This is one of the reasons why registered dietitians are increasingly sought after for advice. Registered dietitians help plan nutrition programs for individuals in various settings. They often work with patients in health care facilities, designing nutrition plans to treat and prevent diseases. For example, dietitians may teach a patient with diabetes how to manage blood sugar levels by eating the correct types and amounts of carbohydrates. Dietitians may also work in nursing homes, schools, and private practices. To become a registered dietitian, one needs to earn at least a bachelor's degree in dietetics, nutrition, food technology, or a related field. In addition, registered dietitians must complete a supervised internship

program and pass a national exam. Those who pursue careers in dietetics take courses in nutrition, chemistry, biochemistry, biology, microbiology, and human physiology. Dietitians must become experts in the chemistry and physiology (biological functions) of food (proteins, carbohydrates, and fats).

## **Benefits of Carbohydrates**

Are carbohydrates good for you? People who wish to lose weight are often told that carbohydrates are bad for them and should be avoided. Some diets completely forbid carbohydrate consumption, claiming that a low-carbohydrate diet helps people to lose weight faster. However, carbohydrates have been an important part of the human diet for thousands of years; artifacts from ancient civilizations show the presence of wheat, rice, and corn in our ancestors' storage areas.

Carbohydrates should be supplemented with proteins, vitamins, and fats to be parts of a well-balanced diet. Calorie-wise, a gram of carbohydrate provides 4.3 Kcal. For comparison, fats provide 9 Kcal/g, a less desirable ratio. Carbohydrates contain soluble and insoluble elements; the insoluble part is known as fiber, which is mostly cellulose. Fiber has many uses; it promotes regular bowel movement by adding bulk, and it regulates the rate of consumption of blood glucose. Fiber also helps to remove excess cholesterol from the body: fiber binds to the cholesterol in the small intestine, then attaches to the cholesterol and prevents the cholesterol particles from entering the bloodstream, and then cholesterol exits the body via the feces. Fiber-rich diets also have a protective role in reducing the occurrence of colon cancer. In addition, a meal containing whole grains and vegetables gives a feeling of fullness. As an immediate source of energy, glucose is broken down during the process of cellular respiration, which produces ATP, the energy currency of the cell. Without the consumption of carbohydrates, the availability of "instant energy" would be reduced. Eliminating carbohydrates from the diet is not the best way to lose weight. A low-calorie diet that is rich in whole grains, fruits, vegetables, and lean

meat, together with plenty of exercise and plenty of water, is the more sensible way to lose weight.

**Note:**

Link to Learning



For an additional perspective on carbohydrates, explore “Biomolecules: the Carbohydrates” through this [interactive animation](#).

## Section Summary

Carbohydrates are a group of macromolecules that are a vital energy source for the cell and provide structural support to plant cells, fungi, and all of the arthropods that include lobsters, crabs, shrimp, insects, and spiders. Carbohydrates are classified as monosaccharides, disaccharides, and polysaccharides depending on the number of monomers in the molecule. Monosaccharides are linked by glycosidic bonds that are formed as a result of dehydration reactions, forming disaccharides and polysaccharides with the elimination of a water molecule for each bond formed. Glucose, galactose, and fructose are common monosaccharides, whereas common disaccharides include lactose, maltose, and sucrose. Starch and glycogen, examples of polysaccharides, are the storage forms of glucose in plants and animals, respectively. The long polysaccharide chains may be branched or unbranched. Cellulose is an example of an unbranched polysaccharide, whereas amylopectin, a constituent of starch, is a highly branched molecule. Storage of glucose, in the form of polymers like starch or glycogen, makes it slightly less accessible for metabolism; however, this prevents it from

leaking out of the cell or creating a high osmotic pressure that could cause excessive water uptake by the cell.

## Art Connections

### Exercise:

**Problem:** [\[link\]](#) What kind of sugars are these, aldose or ketose?

---

### Solution:

[\[link\]](#) Glucose and galactose are aldoses. Fructose is a ketose.

## Review Questions

### Exercise:

**Problem:** An example of a monosaccharide is \_\_\_\_\_.

- a. fructose
- b. glucose
- c. galactose
- d. all of the above

---

### Solution:

D

### Exercise:

**Problem:** Cellulose and starch are examples of:

- a. monosaccharides
- b. disaccharides
- c. lipids



d. polysaccharides

---

**Solution:**

D

**Exercise:**

**Problem:**

Plant cell walls contain which of the following in abundance?

- a. starch
- b. cellulose
- c. glycogen
- d. lactose

---

**Solution:**

B

**Exercise:**

**Problem:**

Lactose is a disaccharide formed by the formation of a \_\_\_\_\_ bond between glucose and \_\_\_\_\_.

- a. glycosidic; lactose
- b. glycosidic; galactose
- c. hydrogen; sucrose
- d. hydrogen; fructose

---

**Solution:**

B

## Free Response

### Exercise:

#### Problem:

Describe the similarities and differences between glycogen and starch.

---

#### Solution:

Glycogen and starch are polysaccharides. They are the storage form of glucose. Glycogen is stored in animals in the liver and in muscle cells, whereas starch is stored in the roots, seeds, and leaves of plants. Starch has two different forms, one unbranched (amylose) and one branched (amylopectin), whereas glycogen is a single type of a highly branched molecule.

### Exercise:

#### Problem:

Why is it impossible for humans to digest food that contains cellulose?

---

#### Solution:

The  $\beta$  1-4 glycosidic linkage in cellulose cannot be broken down by human digestive enzymes. Herbivores such as cows, koalas, and buffalos are able to digest grass that is rich in cellulose and use it as a food source because bacteria and protists in their digestive systems, especially in the rumen, secrete the enzyme cellulase. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal.

## Glossary

### carbohydrate

biological macromolecule in which the ratio of carbon to hydrogen and to oxygen is 1:2:1; carbohydrates serve as energy sources and

structural support in cells and form the a cellular exoskeleton of arthropods

cellulose

polysaccharide that makes up the cell wall of plants; provides structural support to the cell

chitin

type of carbohydrate that forms the outer skeleton of all arthropods that include crustaceans and insects; it also forms the cell walls of fungi

disaccharide

two sugar monomers that are linked together by a glycosidic bond

glycogen

storage carbohydrate in animals

glycosidic bond

bond formed by a dehydration reaction between two monosaccharides with the elimination of a water molecule

monosaccharide

single unit or monomer of carbohydrates

polysaccharide

long chain of monosaccharides; may be branched or unbranched

starch

storage carbohydrate in plants

## Lipids

By the end of this section, you will be able to:

- Describe the four major types of lipids
- Explain the role of fats in storing energy
- Differentiate between saturated and unsaturated fatty acids
- Describe phospholipids and their role in cells
- Define the basic structure of a steroid and some functions of steroids
- Explain the how cholesterol helps to maintain the fluid nature of the plasma membrane

**Lipids** include a diverse group of compounds that are largely nonpolar in nature. This is because they are hydrocarbons that include mostly nonpolar carbon–carbon or carbon–hydrogen bonds. Non-polar molecules are hydrophobic (“water fearing”), or insoluble in water. Lipids perform many different functions in a cell. Cells store energy for long-term use in the form of fats. Lipids also provide insulation from the environment for plants and animals ([\[link\]](#)). For example, they help keep aquatic birds and mammals dry when forming a protective layer over fur or feathers because of their water-repellant hydrophobic nature. Lipids are also the building blocks of many hormones and are an important constituent of all cellular membranes. Lipids include fats, oils, waxes, phospholipids, and steroids.



Hydrophobic lipids in the fur  
of aquatic mammals, such as

this river otter, protect them  
from the elements. (credit:  
Ken Bosma)

## **Fats and Oils**

A fat molecule consists of two main components—glycerol and fatty acids. Glycerol is an organic compound (alcohol) with three carbons, five hydrogens, and three hydroxyl (OH) groups. Fatty acids have a long chain of hydrocarbons to which a carboxyl group is attached, hence the name “fatty acid.” The number of carbons in the fatty acid may range from 4 to 36; most common are those containing 12–18 carbons. In a fat molecule, the fatty acids are attached to each of the three carbons of the glycerol molecule with an ester bond through an oxygen atom ([\[link\]](#)).

$$\begin{array}{c} \text{H} \\ | \\ \text{H}-\text{C}-\text{OH} \\ | \\ \text{H}-\text{C}-\text{OH} \\ | \\ \text{H}-\text{C}-\text{OH} \\ | \\ \text{H} \end{array}$$

+

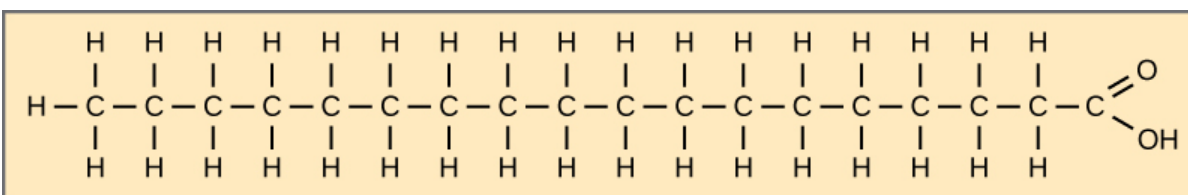
CCCCCCCCCCCCCCCCO

The diagram illustrates three parallel chemical structures of fatty acids, each consisting of a carboxyl group (H-C(=O)-O-) linked to a hydrocarbon chain. The top chain has 10 carbons, the middle has 12, and the bottom has 14. All chains are saturated.

During this ester bond formation, three water molecules are released. The three fatty acids in the triacylglycerol may be similar or dissimilar. Fats are also called **triacylglycerols** or **triglycerides** because of their chemical

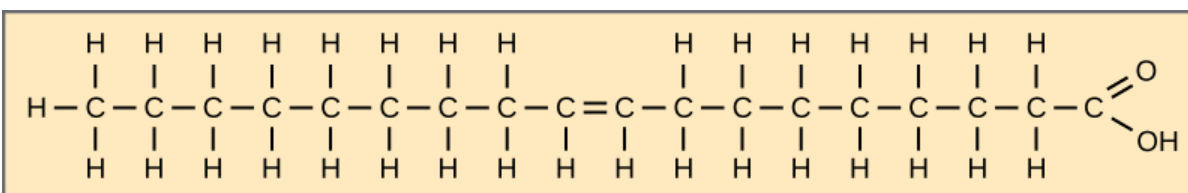
structure. Some fatty acids have common names that specify their origin. For example, palmitic acid, a **saturated fatty acid**, is derived from the palm tree. Arachidic acid is derived from *Arachis hypogea*, the scientific name for groundnuts or peanuts.

Fatty acids may be saturated or unsaturated. In a fatty acid chain, if there are only single bonds between neighboring carbons in the hydrocarbon chain, the fatty acid is said to be saturated. Saturated fatty acids are saturated with hydrogen; in other words, the number of hydrogen atoms attached to the carbon skeleton is maximized. Stearic acid is an example of a saturated fatty acid ([\[link\]](#))



Stearic acid is a common saturated fatty acid.

When the hydrocarbon chain contains a double bond, the fatty acid is said to be **unsaturated**. Oleic acid is an example of an unsaturated fatty acid ([\[link\]](#)).



Oleic acid is a common unsaturated fatty acid.

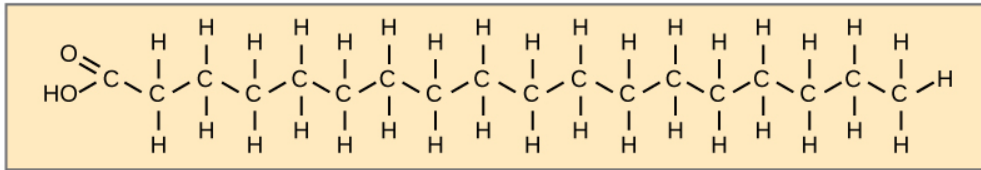
Most unsaturated fats are liquid at room temperature and are called oils. If there is one double bond in the molecule, then it is known as a monounsaturated fat (e.g., olive oil), and if there is more than one double bond, then it is known as a polyunsaturated fat (e.g., canola oil).

When a fatty acid has no double bonds, it is known as a saturated fatty acid because no more hydrogen may be added to the carbon atoms of the chain. A fat may contain similar or different fatty acids attached to glycerol. Long straight fatty acids with single bonds tend to get packed tightly and are solid at room temperature. Animal fats with stearic acid and palmitic acid (common in meat) and the fat with butyric acid (common in butter) are examples of saturated fats. Mammals store fats in specialized cells called adipocytes, where globules of fat occupy most of the cell's volume. In plants, fat or oil is stored in many seeds and is used as a source of energy during seedling development. Unsaturated fats or oils are usually of plant origin and contain *cis* unsaturated fatty acids. *Cis* and *trans* indicate the configuration of the molecule around the double bond. If hydrogens are present in the same plane, it is referred to as a *cis* fat; if the hydrogen atoms are on two different planes, it is referred to as a **trans fat**. The *cis* double bond causes a bend or a “kink” that prevents the fatty acids from packing tightly, keeping them liquid at room temperature ([\[link\]](#)). Olive oil, corn oil, canola oil, and cod liver oil are examples of unsaturated fats. Unsaturated fats help to lower blood cholesterol levels whereas saturated fats contribute to plaque formation in the arteries.



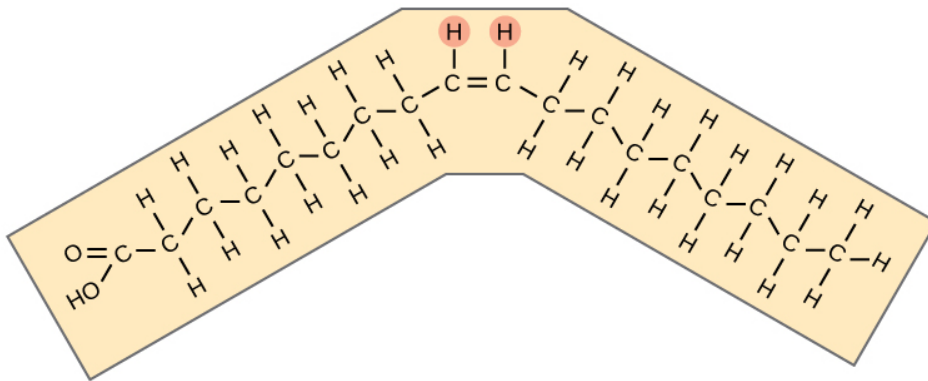
## Saturated fatty acid

Stearic acid

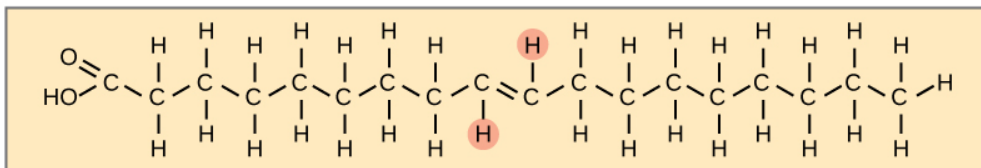


## Unsaturated fatty acids

*Cis* oleic acid



*Trans* oleic acid



Saturated fatty acids have hydrocarbon chains connected by single bonds only. Unsaturated fatty acids have one or more double bonds. Each double bond may be in a *cis* or *trans* configuration. In the *cis* configuration, both hydrogens are on the same side of the hydrocarbon chain. In the *trans* configuration, the hydrogens are on opposite sides. A *cis* double bond causes a kink in the chain.

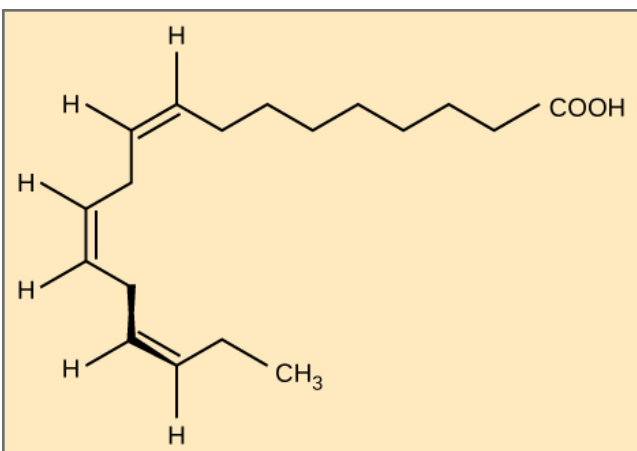
## Trans Fats

In the food industry, oils are artificially hydrogenated to make them semi-solid and of a consistency desirable for many processed food products. Simply speaking, hydrogen gas is bubbled through oils to solidify them. During this hydrogenation process, double bonds of the *cis*- conformation in the hydrocarbon chain may be converted to double bonds in the *trans*-conformation.

Margarine, some types of peanut butter, and shortening are examples of artificially hydrogenated *trans* fats. Recent studies have shown that an increase in *trans* fats in the human diet may lead to an increase in levels of low-density lipoproteins (LDL), or “bad” cholesterol, which in turn may lead to plaque deposition in the arteries, resulting in heart disease. Many fast food restaurants have recently banned the use of *trans* fats, and food labels are required to display the *trans* fat content.

## **Omega Fatty Acids**

Essential fatty acids are fatty acids required but not synthesized by the human body. Consequently, they have to be supplemented through ingestion via the diet. **Omega-3** fatty acids (like that shown in [\[link\]](#)) fall into this category and are one of only two known for humans (the other being omega-6 fatty acid). These are polyunsaturated fatty acids and are called omega-3 because the third carbon from the end of the hydrocarbon chain is connected to its neighboring carbon by a double bond.



Alpha-linolenic acid is an example of an omega-3 fatty acid. It has three *cis* double bonds and, as a result, a curved shape. For clarity, the carbons are not shown. Each singly bonded carbon has two hydrogens associated with it, also not shown.

The farthest carbon away from the carboxyl group is numbered as the omega ( $\omega$ ) carbon, and if the double bond is between the third and fourth carbon from that end, it is known as an omega-3 fatty acid. Nutritionally important because the body does not make them, omega-3 fatty acids include alpha-linolenic acid (ALA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA), all of which are polyunsaturated. Salmon, trout, and tuna are good sources of omega-3 fatty acids. Research indicates that omega-3 fatty acids reduce the risk of sudden death from heart attacks, reduce triglycerides in the blood, lower blood pressure, and prevent thrombosis by inhibiting blood clotting. They also reduce inflammation, and may help reduce the risk of some cancers in animals.

Like carbohydrates, fats have received a lot of bad publicity. It is true that eating an excess of fried foods and other “fatty” foods leads to weight gain. However, fats do have important functions. Many vitamins are fat soluble,

and fats serve as a long-term storage form of fatty acids: a source of energy. They also provide insulation for the body. Therefore, “healthy” fats in moderate amounts should be consumed on a regular basis.

## Waxes

**Wax** covers the feathers of some aquatic birds and the leaf surfaces of some plants. Because of the hydrophobic nature of waxes, they prevent water from sticking on the surface ([link](#)). Waxes are made up of long fatty acid chains esterified to long-chain alcohols.

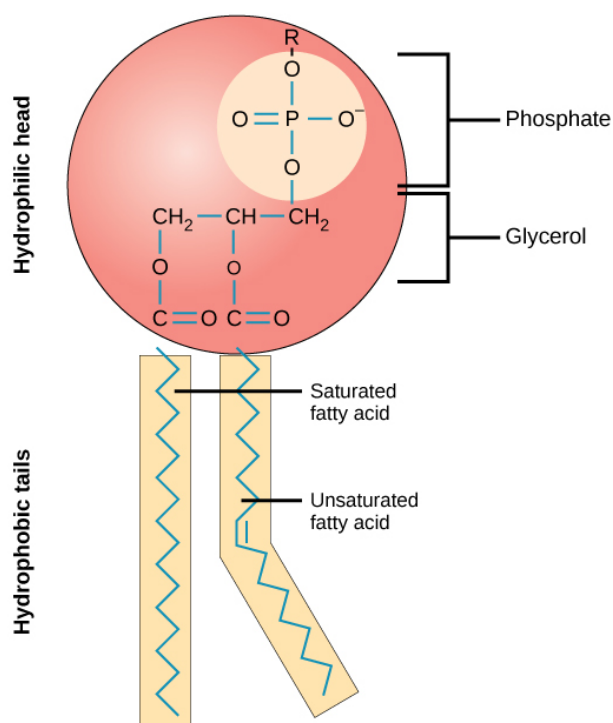


Waxy coverings on some leaves are made of lipids. (credit: Roger Griffith)

## Phospholipids

**Phospholipids** are major constituents of the plasma membrane, the outermost layer of animal cells. Like fats, they are composed of fatty acid chains attached to a glycerol or sphingosine backbone. Instead of three fatty acids attached as in triglycerides, however, there are two fatty acids forming

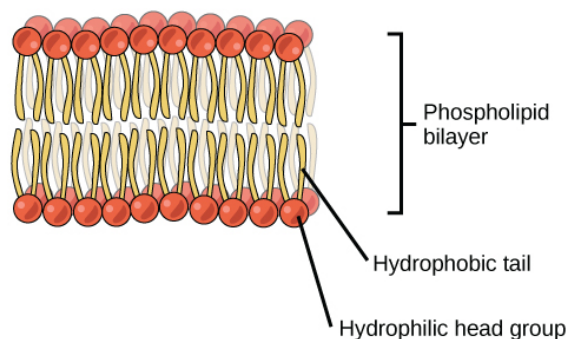
diacylglycerol, and the third carbon of the glycerol backbone is occupied by a modified phosphate group ([\[link\]](#)). A phosphate group alone attached to a diacylglycerol does not qualify as a phospholipid; it is phosphatidate (diacylglycerol 3-phosphate), the precursor of phospholipids. The phosphate group is modified by an alcohol. Phosphatidylcholine and phosphatidylserine are two important phospholipids that are found in plasma membranes.



A phospholipid is a molecule with two fatty acids and a modified phosphate group attached to a glycerol backbone. The phosphate may be modified by the addition of charged or polar chemical groups. Two chemical groups that may modify the phosphate, choline and serine, are shown here. Both choline and serine attach to the phosphate group at

the position labeled R via the hydroxyl group indicated in green.

A phospholipid is an amphipathic molecule, meaning it has a hydrophobic and a hydrophilic part. The fatty acid chains are hydrophobic and cannot interact with water, whereas the phosphate-containing group is hydrophilic and interacts with water ([\[link\]](#)).



The phospholipid bilayer is the major component of all cellular membranes. The hydrophilic head groups of the phospholipids face the aqueous solution. The hydrophobic tails are sequestered in the middle of the bilayer.

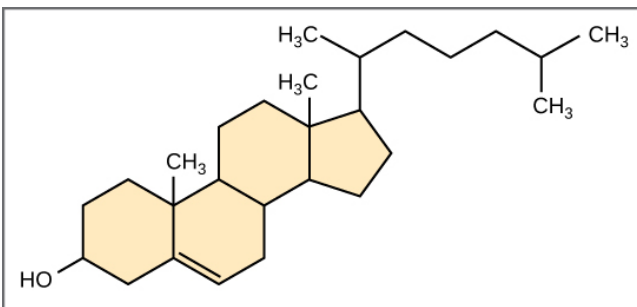
The head is the hydrophilic part, and the tail contains the hydrophobic fatty acids. In a membrane, a bilayer of phospholipids forms the matrix of the structure, the fatty acid tails of phospholipids face inside, away from water, whereas the phosphate group faces the outside, aqueous side ([\[link\]](#)).

Phospholipids are responsible for the dynamic nature of the plasma membrane. If a drop of phospholipids is placed in water, it spontaneously

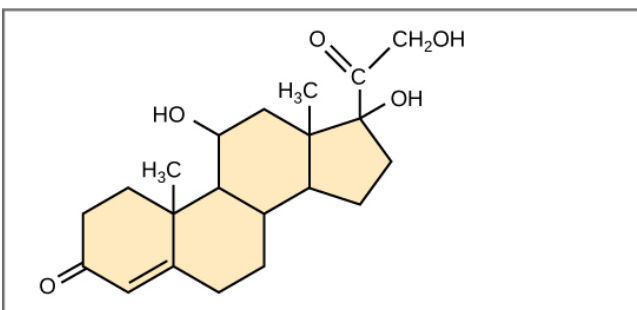
forms a structure known as a micelle, where the hydrophilic phosphate heads face the outside and the fatty acids face the interior of this structure.

## Steroids

Unlike the phospholipids and fats discussed earlier, **steroids** have a fused ring structure. Although they do not resemble the other lipids, they are grouped with them because they are also hydrophobic and insoluble in water. All steroids have four linked carbon rings and several of them, like cholesterol, have a short tail ([link](#)). Many steroids also have the  $\text{-OH}$  functional group, which puts them in the alcohol classification (sterols).



**Cholesterol**



**Cortisol**

Steroids such as cholesterol and cortisol are composed of four fused hydrocarbon rings.

Cholesterol is the most common steroid. Cholesterol is mainly synthesized in the liver and is the precursor to many steroid hormones such as testosterone and estradiol, which are secreted by the gonads and endocrine glands. It is also the precursor to Vitamin D. Cholesterol is also the precursor of bile salts, which help in the emulsification of fats and their subsequent absorption by cells. Although cholesterol is often spoken of in negative terms by lay people, it is necessary for proper functioning of the body. It is a component of the plasma membrane of animal cells and is found within the phospholipid bilayer. Being the outermost structure in animal cells, the plasma membrane is responsible for the transport of materials and cellular recognition and it is involved in cell-to-cell communication.

**Note:**

Link to Learning



For an additional perspective on lipids, explore the interactive animation [“Biomolecules: The Lipids”](#).

## Section Summary

Lipids are a class of macromolecules that are nonpolar and hydrophobic in nature. Major types include fats and oils, waxes, phospholipids, and steroids. Fats are a stored form of energy and are also known as triacylglycerols or triglycerides. Fats are made up of fatty acids and either glycerol or sphingosine. Fatty acids may be unsaturated or saturated,



depending on the presence or absence of double bonds in the hydrocarbon chain. If only single bonds are present, they are known as saturated fatty acids. Unsaturated fatty acids may have one or more double bonds in the hydrocarbon chain. Phospholipids make up the matrix of membranes. They have a glycerol or sphingosine backbone to which two fatty acid chains and a phosphate-containing group are attached. Steroids are another class of lipids. Their basic structure has four fused carbon rings. Cholesterol is a type of steroid and is an important constituent of the plasma membrane, where it helps to maintain the fluid nature of the membrane. It is also the precursor of steroid hormones such as testosterone.

## Review Questions

### Exercise:

#### Problem:

Saturated fats have all of the following characteristics except:

- a. they are solid at room temperature
- b. they have single bonds within the carbon chain
- c. they are usually obtained from animal sources
- d. they tend to dissolve in water easily

---

#### Solution:

D

### Exercise:

**Problem:** Phospholipids are important components of \_\_\_\_\_.

- a. the plasma membrane of animal cells
  - b. the ring structure of steroids
  - c. the waxy covering on leaves
  - d. the double bond in hydrocarbon chains
-

**Solution:**

A

**Free Response****Exercise:****Problem:**

Explain at least three functions that lipids serve in plants and/or animals.

---

**Solution:**

Fat serves as a valuable way for animals to store energy. It can also provide insulation. Waxes can protect plant leaves and mammalian fur from getting wet. Phospholipids and steroids are important components of animal cell membranes, as well as plant, fungal, and bacterial membranes.

**Exercise:****Problem:**

Why have trans fats been banned from some restaurants? How are they created?

---

**Solution:**

Trans fats are created artificially when hydrogen gas is bubbled through oils to solidify them. The double bonds of the *cis* conformation in the hydrocarbon chain may be converted to double bonds in the *trans* configuration. Some restaurants are banning trans fats because they cause higher levels of LDL, or “bad” cholesterol.

**Glossary**

lipid

macromolecule that is nonpolar and insoluble in water

omega fat

type of polyunsaturated fat that is required by the body; the numbering of the carbon omega starts from the methyl end or the end that is farthest from the carboxylic end

phospholipid

major constituent of the membranes; composed of two fatty acids and a phosphate-containing group attached to a glycerol backbone

saturated fatty acid

long-chain of hydrocarbon with single covalent bonds in the carbon chain; the number of hydrogen atoms attached to the carbon skeleton is maximized

steroid

type of lipid composed of four fused hydrocarbon rings forming a planar structure

trans fat

fat formed artificially by hydrogenating oils, leading to a different arrangement of double bond(s) than those found in naturally occurring lipids

triacylglycerol (also, triglyceride)

fat molecule; consists of three fatty acids linked to a glycerol molecule

unsaturated fatty acid

long-chain hydrocarbon that has one or more double bonds in the hydrocarbon chain

wax

lipid made of a long-chain fatty acid that is esterified to a long-chain alcohol; serves as a protective coating on some feathers, aquatic mammal fur, and leaves

## Proteins

By the end of this section, you will be able to:

- Describe the functions proteins perform in the cell and in tissues
- Discuss the relationship between amino acids and proteins
- Explain the four levels of protein organization
- Describe the ways in which protein shape and function are linked

**Proteins** are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules.

Proteins may be structural, regulatory, contractile, or protective; they may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, polymers of amino acids, arranged in a linear sequence.

## Types and Functions of Proteins

**Enzymes**, which are produced by living cells, are catalysts in biochemical reactions (like digestion) and are usually complex or conjugated proteins. Each enzyme is specific for the substrate (a reactant that binds to an enzyme) it acts on. The enzyme may help in breakdown, rearrangement, or synthesis reactions. Enzymes that break down their substrates are called catabolic enzymes, enzymes that build more complex molecules from their substrates are called anabolic enzymes, and enzymes that affect the rate of reaction are called catalytic enzymes. It should be noted that all enzymes increase the rate of reaction and, therefore, are considered to be organic catalysts. An example of an enzyme is salivary amylase, which hydrolyzes its substrate amylose, a component of starch.

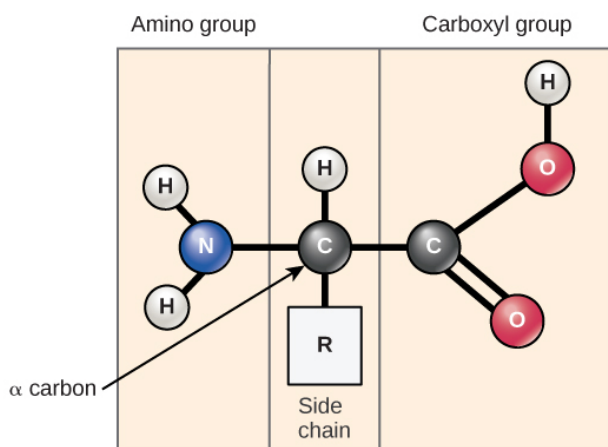
**Hormones** are chemical-signaling molecules, usually small proteins or steroids, secreted by endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that helps to regulate the blood glucose level. The primary types and functions of proteins are listed in [\[link\]](#).

Protein Types and Functions		
Type	Examples	Functions
Digestive Enzymes	Amylase, lipase, pepsin, trypsin	Help in digestion of food by catabolizing nutrients into monomeric units
Transport	Hemoglobin, albumin	Carry substances in the blood or lymph throughout the body
Structural	Actin, tubulin, keratin	Construct different structures, like the cytoskeleton
Hormones	Insulin, thyroxine	Coordinate the activity of different body systems
Defense	Immunoglobulins	Protect the body from foreign pathogens
Contractile	Actin, myosin	Effect muscle contraction
Storage	Legume storage proteins, egg white (albumin)	Provide nourishment in early development of the embryo and the seedling

Proteins have different shapes and molecular weights; some proteins are globular in shape whereas others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, found in our skin, is a fibrous protein. Protein shape is critical to its function, and this shape is maintained by many different types of chemical bonds. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the shape of the protein, leading to loss of function, known as **denaturation**. All proteins are made up of different arrangements of the same 20 types of amino acids.

## Amino Acids

**Amino acids** are the monomers that make up proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom, also known as the alpha ( $\alpha$ ) carbon, bonded to an amino group ( $\text{NH}_2$ ), a carboxyl group ( $\text{COOH}$ ), and to a hydrogen atom. Every amino acid also has another atom or group of atoms bonded to the central atom known as the R group ([\[link\]](#)).



Amino acids have a central asymmetric carbon to which an amino group, a carboxyl group, a hydrogen atom, and a side chain (R group) are attached.

The name "amino acid" is derived from the fact that they contain both amino group and carboxyl-acid-group in their basic structure. As mentioned, there are 20 amino acids present in proteins. Nine of these are considered essential amino acids in humans because the human body cannot produce them and they are obtained from the diet. For each amino acid, the R group (or side chain) is different ([\[link\]](#)).

**Note:****Art Connection**

AMINO ACID			
Nonpolar, aliphatic R groups			
	Glycine	Alanine	Valine
	Leucine	Methionine	Isoleucine
Polar, uncharged R groups			
	Serine	Threonine	Cysteine
	Proline	Asparagine	Glutamine
AMINO ACID			
Positively charged R groups			
	Lysine	Arginine	Histidine
Negatively charged R groups			
	Aspartate	Glutamate	
Nonpolar, aromatic R groups			
	Phenylalanine	Tyrosine	Tryptophan

There are 20 common amino acids commonly found in proteins, each with a different R group (variant group) that determines its chemical nature.

Which categories of amino acid would you expect to find on the surface of a soluble protein, and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

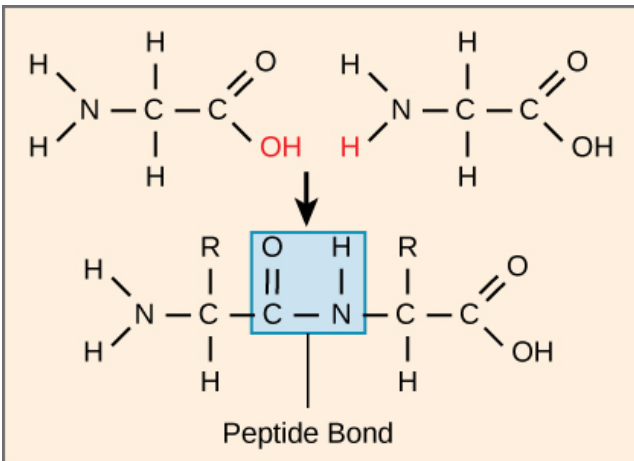
The chemical nature of the side chain determines the nature of the amino acid (that is, whether it is acidic, basic, polar, or nonpolar). For example, the

amino acid glycine has a hydrogen atom as the R group. Amino acids such as valine, methionine, and alanine are nonpolar or hydrophobic in nature, while amino acids such as serine, threonine, and cysteine are polar and have hydrophilic side chains. The side chains of lysine and arginine are positively charged, and therefore these amino acids are also known as basic amino acids. Proline has an R group that is linked to the amino group, forming a ring-like structure. Proline is an exception to the standard structure of an amino acid since its amino group is not separate from the side chain ([\[link\]](#)).

Amino acids are represented by a single upper case letter or a three-letter abbreviation. For example, valine is known by the letter V or the three-letter symbol val. Just as some fatty acids are essential to a diet, some amino acids are necessary as well. They are known as essential amino acids, and in humans they include isoleucine, leucine, and cysteine. Essential amino acids refer to those necessary for construction of proteins in the body, although not produced by the body; which amino acids are essential varies from organism to organism.

The sequence and the number of amino acids ultimately determine the protein's shape, size, and function. Each amino acid is attached to another amino acid by a covalent bond, known as a **peptide bond**, which is formed by a dehydration reaction. The carboxyl group of one amino acid and the amino group of the incoming amino acid combine, releasing a molecule of water. The resulting bond is the peptide bond ([\[link\]](#)).





Peptide bond formation is a dehydration synthesis reaction. The carboxyl group of one amino acid is linked to the amino group of the incoming amino acid. In the process, a molecule of water is released.

The products formed by such linkages are called peptides. As more amino acids join to this growing chain, the resulting chain is known as a polypeptide. Each polypeptide has a free amino group at one end. This end is called the N terminal, or the amino terminal, and the other end has a free carboxyl group, also known as the C or carboxyl terminal. While the terms polypeptide and protein are sometimes used interchangeably, a polypeptide is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, often have bound non-peptide prosthetic groups, have a distinct shape, and have a unique function. After protein synthesis (translation), most proteins are modified. These are known as post-translational modifications. They may undergo cleavage, phosphorylation, or may require the addition of other chemical groups. Only after these modifications is the protein completely functional.

**Note:**

Link to Learning



Click through the steps of protein synthesis in this [interactive tutorial](#).

**Note:**

Evolution Connection

**The Evolutionary Significance of Cytochrome c**

Cytochrome c is an important component of the electron transport chain, a part of cellular respiration, and it is normally found in the cellular organelle, the mitochondrion. This protein has a heme prosthetic group, and the central ion of the heme gets alternately reduced and oxidized during electron transfer. Because this essential protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of cytochrome c amino acid sequence homology among different species; in other words, evolutionary kinship can be assessed by measuring the similarities or differences among various species' DNA or protein sequences.

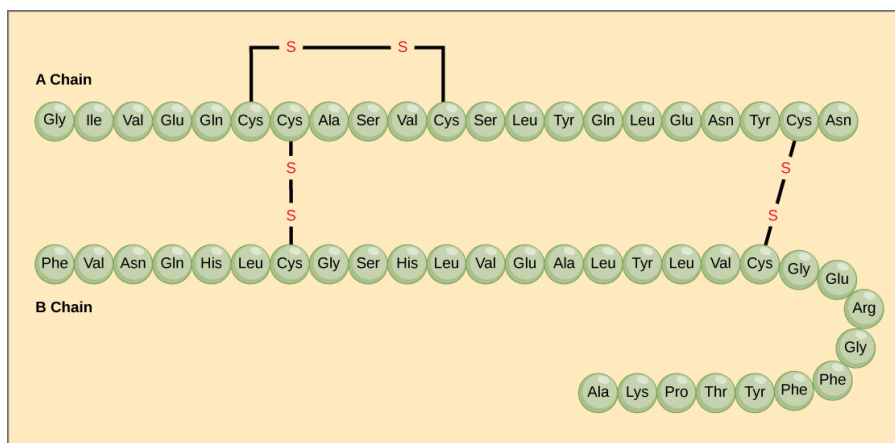
Scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule from different organisms that has been sequenced to date, 37 of these amino acids appear in the same position in all samples of cytochrome c. This indicates that there may have been a common ancestor. On comparing the human and chimpanzee protein sequences, no sequence difference was found. When human and rhesus monkey sequences were compared, the single difference found was in one amino acid. In another comparison, human to yeast sequencing shows a difference in the 44th position.

## Protein Structure

As discussed earlier, the shape of a protein is critical to its function. For example, an enzyme can bind to a specific substrate at a site known as the active site. If this active site is altered because of local changes or changes in overall protein structure, the enzyme may be unable to bind to the substrate. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary.

### Primary Structure

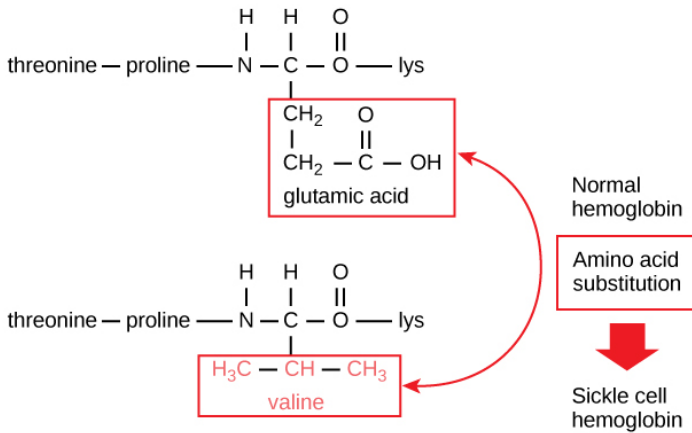
The unique sequence of amino acids in a polypeptide chain is its **primary structure**. For example, the pancreatic hormone insulin has two polypeptide chains, A and B, and they are linked together by disulfide bonds. The N terminal amino acid of the A chain is glycine, whereas the C terminal amino acid is asparagine ([\[link\]](#)). The sequences of amino acids in the A and B chains are unique to insulin.



Bovine serum insulin is a protein hormone made of two peptide chains, A (21 amino acids long) and B (30 amino acids long). In each chain, primary structure is indicated by three-letter abbreviations

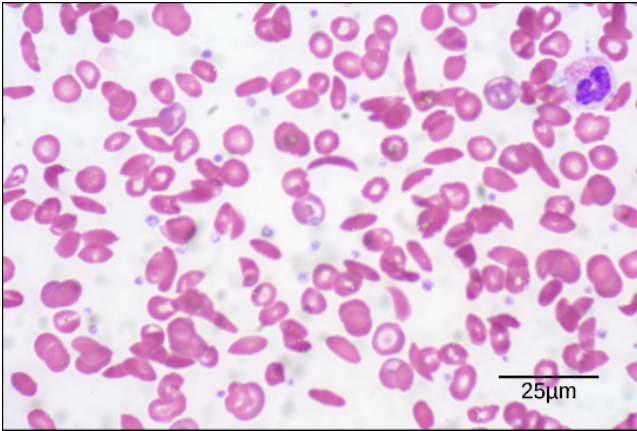
that represent the names of the amino acids in the order they are present. The amino acid cysteine (cys) has a sulfhydryl (SH) group as a side chain. Two sulfhydryl groups can react in the presence of oxygen to form a disulfide (S-S) bond. Two disulfide bonds connect the A and B chains together, and a third helps the A chain fold into the correct shape. Note that all disulfide bonds are the same length, but are drawn different sizes for clarity.

The unique sequence for every protein is ultimately determined by the gene encoding the protein. A change in nucleotide sequence of the gene's coding region may lead to a different amino acid being added to the growing polypeptide chain, causing a change in protein structure and function. In sickle cell anemia, the hemoglobin  $\beta$  chain (a small portion of which is shown in [\[link\]](#)) has a single amino acid substitution, causing a change in protein structure and function. Specifically, the amino acid glutamic acid is substituted by valine in the  $\beta$  chain. What is most remarkable to consider is that a hemoglobin molecule is made up of two alpha chains and two beta chains that each consist of about 150 amino acids. The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—which dramatically decreases life expectancy—is a single amino acid of the 600. What is even more remarkable is that those 600 amino acids are encoded by three nucleotides each, and the mutation is caused by a single base change (point mutation), 1 in 1800 bases.



The beta chain of hemoglobin is 147 residues in length, yet a single amino acid substitution leads to sickle cell anemia. In normal hemoglobin, the amino acid at position seven is glutamate. In sickle cell hemoglobin, this glutamate is replaced by a valine.

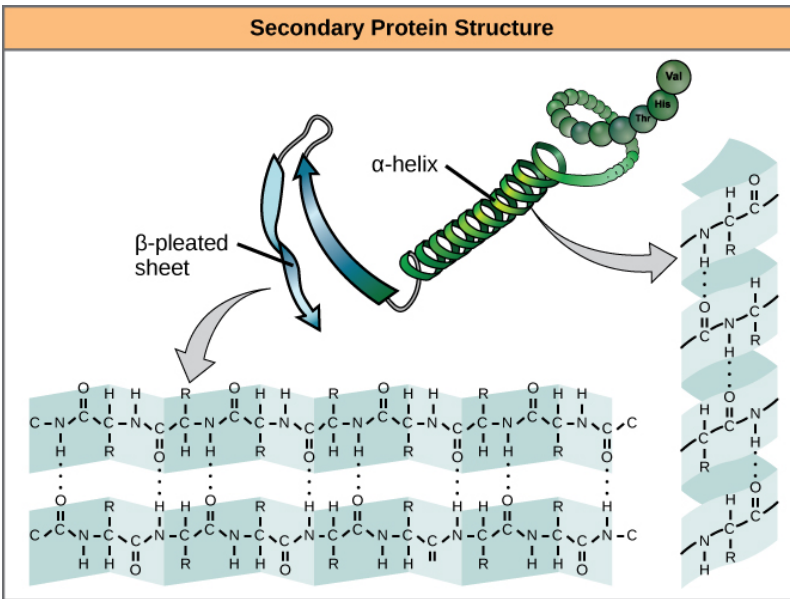
Because of this change of one amino acid in the chain, hemoglobin molecules form long fibers that distort the biconcave, or disc-shaped, red blood cells and assume a crescent or “sickle” shape, which clogs arteries ([link](#)). This can lead to myriad serious health problems such as breathlessness, dizziness, headaches, and abdominal pain for those affected by this disease.



In this blood smear, visualized at 535x magnification using bright field microscopy, sickle cells are crescent shaped, while normal cells are disc-shaped. (credit: modification of work by Ed Uthman; scale-bar data from Matt Russell)

## Secondary Structure

The local folding of the polypeptide in some regions gives rise to the **secondary structure** of the protein. The most common are the  **$\alpha$ -helix** and  **$\beta$ -pleated sheet** structures ([\[link\]](#)). Both structures are the  $\alpha$ -helix structure—the helix held in shape by hydrogen bonds. The hydrogen bonds form between the oxygen atom in the carbonyl group in one amino acid and another amino acid that is four amino acids farther along the chain.

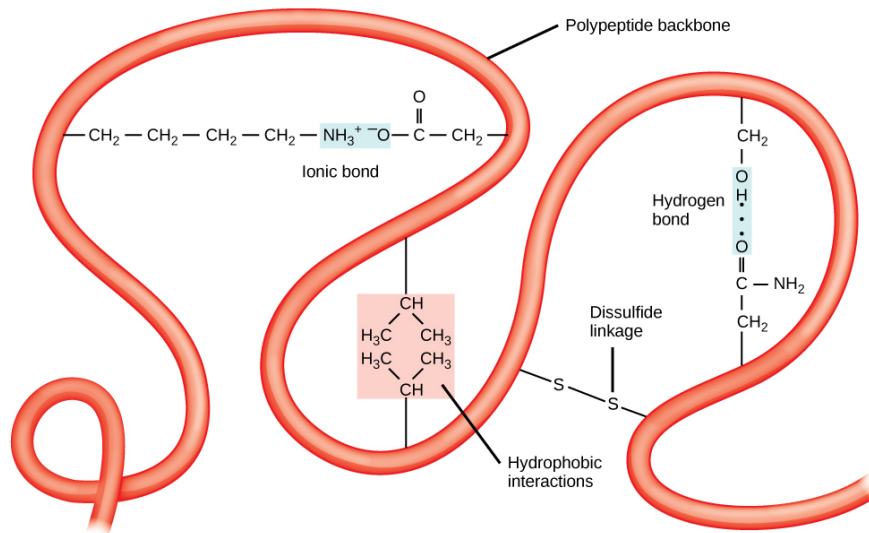


The  $\alpha$ -helix and  $\beta$ -pleated sheet are secondary structures of proteins that form because of hydrogen bonding between carbonyl and amino groups in the peptide backbone. Certain amino acids have a propensity to form an  $\alpha$ -helix, while others have a propensity to form a  $\beta$ -pleated sheet.

Every helical turn in an alpha helix has 3.6 amino acid residues. The R groups (the variant groups) of the polypeptide protrude out from the  $\alpha$ -helix chain. In the  $\beta$ -pleated sheet, the “pleats” are formed by hydrogen bonding between atoms on the backbone of the polypeptide chain. The R groups are attached to the carbons and extend above and below the folds of the pleat. The pleated segments align parallel or antiparallel to each other, and hydrogen bonds form between the partially positive nitrogen atom in the amino group and the partially negative oxygen atom in the carbonyl group of the peptide backbone. The  $\alpha$ -helix and  $\beta$ -pleated sheet structures are found in most globular and fibrous proteins and they play an important structural role.

## Tertiary Structure

The unique three-dimensional structure of a polypeptide is its **tertiary structure** ([\[link\]](#)). This structure is in part due to chemical interactions at work on the polypeptide chain. Primarily, the interactions among R groups creates the complex three-dimensional tertiary structure of a protein. The nature of the R groups found in the amino acids involved can counteract the formation of the hydrogen bonds described for standard secondary structures. For example, R groups with like charges are repelled by each other and those with unlike charges are attracted to each other (ionic bonds). When protein folding takes place, the hydrophobic R groups of nonpolar amino acids lay in the interior of the protein, whereas the hydrophilic R groups lay on the outside. The former types of interactions are also known as hydrophobic interactions. Interaction between cysteine side chains forms disulfide linkages in the presence of oxygen, the only covalent bond forming during protein folding.



The tertiary structure of proteins is determined by a variety of chemical interactions. These include hydrophobic interactions, ionic bonding, hydrogen bonding and disulfide linkages.

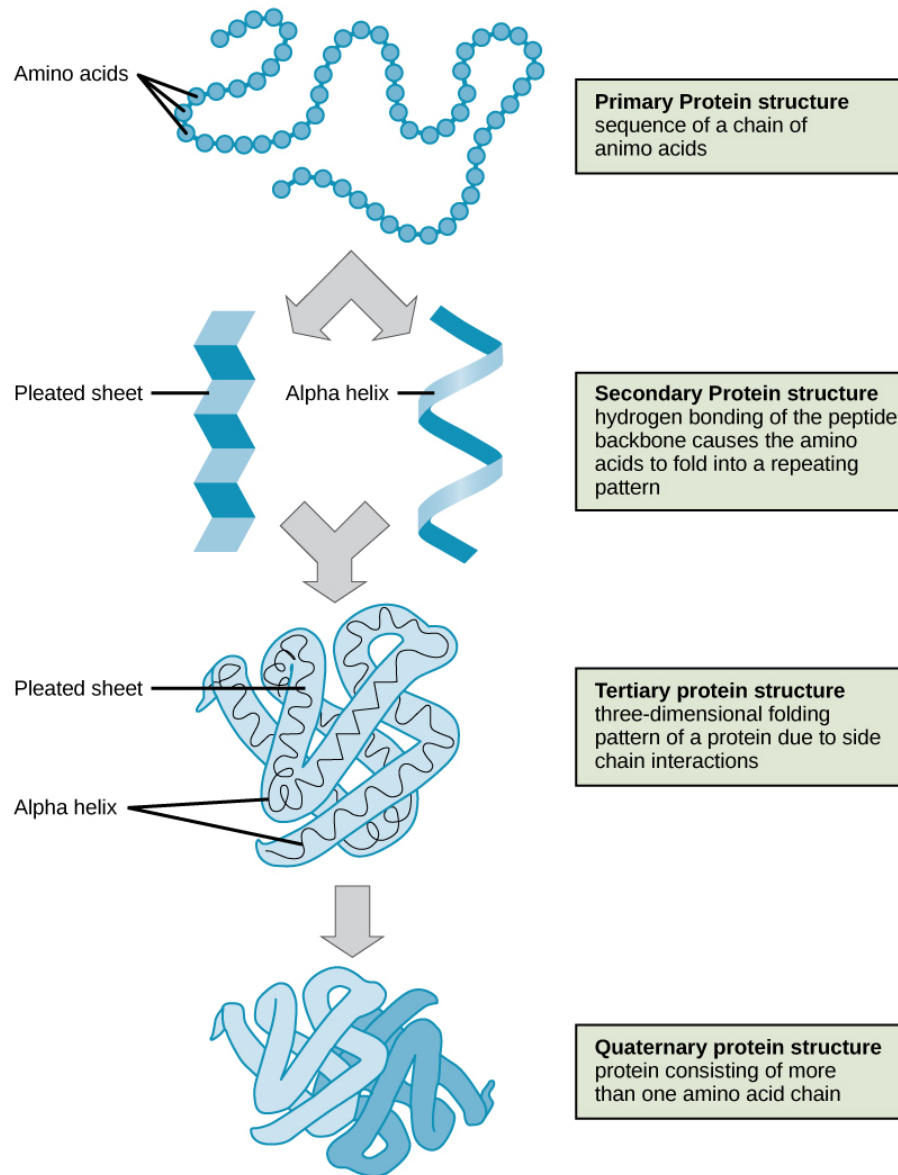


All of these interactions, weak and strong, determine the final three-dimensional shape of the protein. When a protein loses its three-dimensional shape, it may no longer be functional.

## Quaternary Structure

In nature, some proteins are formed from several polypeptides, also known as subunits, and the interaction of these subunits forms the **quaternary structure**. Weak interactions between the subunits help to stabilize the overall structure. For example, insulin (a globular protein) has a combination of hydrogen bonds and disulfide bonds that cause it to be mostly clumped into a ball shape. Insulin starts out as a single polypeptide and loses some internal sequences in the presence of post-translational modification after the formation of the disulfide linkages that hold the remaining chains together. Silk (a fibrous protein), however, has a  $\beta$ -pleated sheet structure that is the result of hydrogen bonding between different chains.

The four levels of protein structure (primary, secondary, tertiary, and quaternary) are illustrated in [\[link\]](#).



The four levels of protein structure can be observed in these illustrations. (credit: modification of work by National Human Genome Research Institute)

## Denaturation and Protein Folding

Each protein has its own unique sequence and shape that are held together by chemical interactions. If the protein is subject to changes in temperature, pH, or exposure to chemicals, the protein structure may change, losing its shape without losing its primary sequence in what is known as denaturation. Denaturation is often reversible because the primary structure of the polypeptide is conserved in the process if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to loss of function. One example of irreversible protein denaturation is when an egg is fried. The albumin protein in the liquid egg white is denatured when placed in a hot pan. Not all proteins are denatured at high temperatures; for instance, bacteria that survive in hot springs have proteins that function at temperatures close to boiling. The stomach is also very acidic, has a low pH, and denatures proteins as part of the digestion process; however, the digestive enzymes of the stomach retain their activity under these conditions.

Protein folding is critical to its function. It was originally thought that the proteins themselves were responsible for the folding process. Only recently was it found that often they receive assistance in the folding process from protein helpers known as **chaperones** (or chaperonins) that associate with the target protein during the folding process. They act by preventing aggregation of polypeptides that make up the complete protein structure, and they disassociate from the protein once the target protein is folded.

**Note:**

Link to Learning



For an additional perspective on proteins, view [this animation](#) called “Biomolecules: The Proteins.”

## Section Summary

Proteins are a class of macromolecules that perform a diverse range of functions for the cell. They help in metabolism by providing structural support and by acting as enzymes, carriers, or hormones. The building blocks of proteins (monomers) are amino acids. Each amino acid has a central carbon that is linked to an amino group, a carboxyl group, a hydrogen atom, and an R group or side chain. There are 20 commonly occurring amino acids, each of which differs in the R group. Each amino acid is linked to its neighbors by a peptide bond. A long chain of amino acids is known as a polypeptide.

Proteins are organized at four levels: primary, secondary, tertiary, and (optional) quaternary. The primary structure is the unique sequence of amino acids. The local folding of the polypeptide to form structures such as the  $\alpha$  helix and  $\beta$ -pleated sheet constitutes the secondary structure. The overall three-dimensional structure is the tertiary structure. When two or more polypeptides combine to form the complete protein structure, the configuration is known as the quaternary structure of a protein. Protein shape and function are intricately linked; any change in shape caused by changes in temperature or pH may lead to protein denaturation and a loss in function.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Which categories of amino acid would you expect to find on the surface of a soluble protein, and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

---

#### Solution:

[\[link\]](#) Polar and charged amino acid residues (the remainder after peptide bond formation) are more likely to be found on the surface of

soluble proteins where they can interact with water, and nonpolar (e.g., amino acid side chains) are more likely to be found in the interior where they are sequestered from water. In membrane proteins, nonpolar and hydrophobic amino acid side chains associate with the hydrophobic tails of phospholipids, while polar and charged amino acid side chains interact with the polar head groups or with the aqueous solution. However, there are exceptions. Sometimes, positively and negatively charged amino acid side chains interact with one another in the interior of a protein, and polar or charged amino acid side chains that interact with a ligand can be found in the ligand binding pocket.

## Review Questions

### Exercise:

**Problem:** The monomers that make up proteins are called \_\_\_\_\_.

- a. nucleotides
- b. disaccharides
- c. amino acids
- d. chaperones

---

### Solution:

C

### Exercise:

#### Problem:

The  $\alpha$  helix and the  $\beta$ -pleated sheet are part of which protein structure?

- a. primary
- b. secondary
- c. tertiary
- d. quaternary

---

**Solution:**

B

**Free Response****Exercise:****Problem:**

Explain what happens if even one amino acid is substituted for another in a polypeptide chain. Provide a specific example.

---

**Solution:**

A change in gene sequence can lead to a different amino acid being added to a polypeptide chain instead of the normal one. This causes a change in protein structure and function. For example, in sickle cell anemia, the hemoglobin  $\beta$  chain has a single amino acid substitution—the amino acid glutamic acid in position six is substituted by valine. Because of this change, hemoglobin molecules form aggregates, and the disc-shaped red blood cells assume a crescent shape, which results in serious health problems.

**Exercise:**

**Problem:** Describe the differences in the four protein structures.

---

**Solution:**

The sequence and number of amino acids in a polypeptide chain is its primary structure. The local folding of the polypeptide in some regions is the secondary structure of the protein. The three-dimensional structure of a polypeptide is known as its tertiary structure, created in part by chemical interactions such as hydrogen bonds between polar side chains, van der Waals interactions, disulfide linkages, and

hydrophobic interactions. Some proteins are formed from multiple polypeptides, also known as subunits, and the interaction of these subunits forms the quaternary structure.

## Glossary

### alpha-helix structure ( $\alpha$ -helix)

type of secondary structure of proteins formed by folding of the polypeptide into a helix shape with hydrogen bonds stabilizing the structure

### amino acid

monomer of a protein; has a central carbon or alpha carbon to which an amino group, a carboxyl group, a hydrogen, and an R group or side chain is attached; the R group is different for all 20 amino acids

### beta-pleated sheet ( $\beta$ -pleated)

secondary structure found in proteins in which “pleats” are formed by hydrogen bonding between atoms on the backbone of the polypeptide chain

### chaperone

(also, chaperonin) protein that helps nascent protein in the folding process

### denaturation

loss of shape in a protein as a result of changes in temperature, pH, or exposure to chemicals

### enzyme

catalyst in a biochemical reaction that is usually a complex or conjugated protein

### hormone

chemical signaling molecule, usually protein or steroid, secreted by endocrine cells that act to control or regulate specific physiological processes

peptide bond

bond formed between two amino acids by a dehydration reaction

polypeptide

long chain of amino acids linked by peptide bonds

primary structure

linear sequence of amino acids in a protein

protein

biological macromolecule composed of one or more chains of amino acids

quaternary structure

association of discrete polypeptide subunits in a protein

secondary structure

regular structure formed by proteins by intramolecular hydrogen bonding between the oxygen atom of one amino acid residue and the hydrogen attached to the nitrogen atom of another amino acid residue

tertiary structure

three-dimensional conformation of a protein, including interactions between secondary structural elements; formed from interactions between amino acid side chains



## Nucleic Acids

By the end of this section, you will be able to:

- Describe the structure of nucleic acids and define the two types of nucleic acids
- Explain the structure and role of DNA
- Explain the structure and roles of RNA

**Nucleic acids** are the most important macromolecules for the continuity of life. They carry the genetic blueprint of a cell and carry instructions for the functioning of the cell.

## DNA and RNA

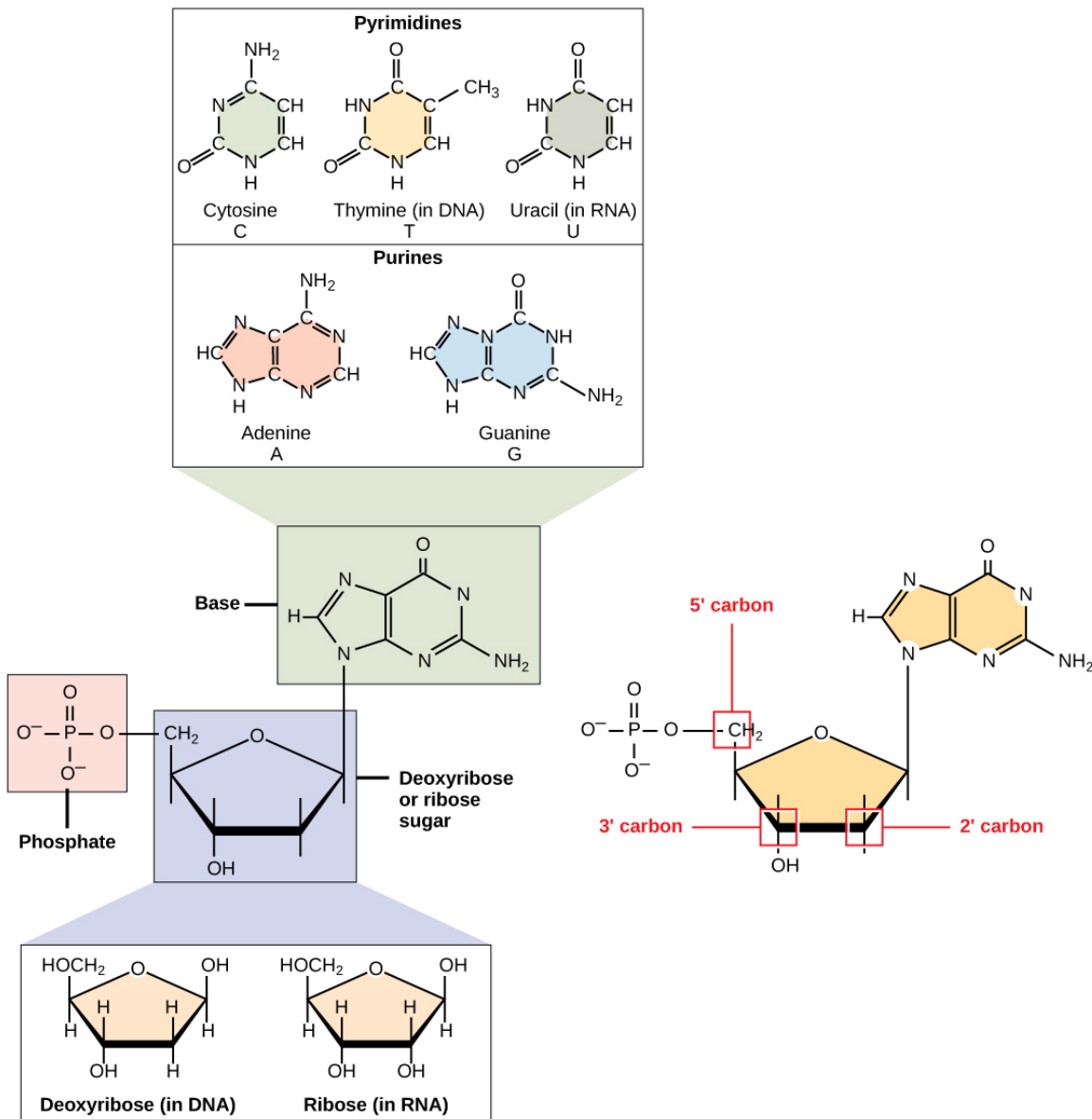
The two main types of nucleic acids are **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA is the genetic material found in all living organisms, ranging from single-celled bacteria to multicellular mammals. It is found in the nucleus of eukaryotes and in the organelles, chloroplasts, and mitochondria. In prokaryotes, the DNA is not enclosed in a membranous envelope.

The entire genetic content of a cell is known as its genome, and the study of genomes is genomics. In eukaryotic cells but not in prokaryotes, DNA forms a complex with histone proteins to form chromatin, the substance of eukaryotic chromosomes. A chromosome may contain tens of thousands of genes. Many genes contain the information to make protein products; other genes code for RNA products. DNA controls all of the cellular activities by turning the genes “on” or “off.”

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus but instead use an intermediary to communicate with the rest of the cell. This intermediary is the **messenger RNA (mRNA)**. Other types of RNA—like rRNA, tRNA, and microRNA—are involved in protein synthesis and its regulation.

DNA and RNA are made up of monomers known as **nucleotides**. The nucleotides combine with each other to form a **polynucleotide**, DNA or

RNA. Each nucleotide is made up of three components: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group ([\[link\]](#)). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.



A nucleotide is made up of three components: a nitrogenous base, a pentose sugar, and one or more phosphate groups. Carbon residues in the pentose are numbered 1' through 5' (the prime distinguishes these residues from those in the base, which are numbered without using a

prime notation). The base is attached to the 1' position of the ribose, and the phosphate is attached to the 5' position. When a polynucleotide is formed, the 5' phosphate of the incoming nucleotide attaches to the 3' hydroxyl group at the end of the growing chain. Two types of pentose are found in nucleotides, deoxyribose (found in DNA) and ribose (found in RNA). Deoxyribose is similar in structure to ribose, but it has an H instead of an OH at the 2' position. Bases can be divided into two categories: purines and pyrimidines. Purines have a double ring structure, and pyrimidines have a single ring.

The nitrogenous bases, important components of nucleotides, are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus, decreases the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).

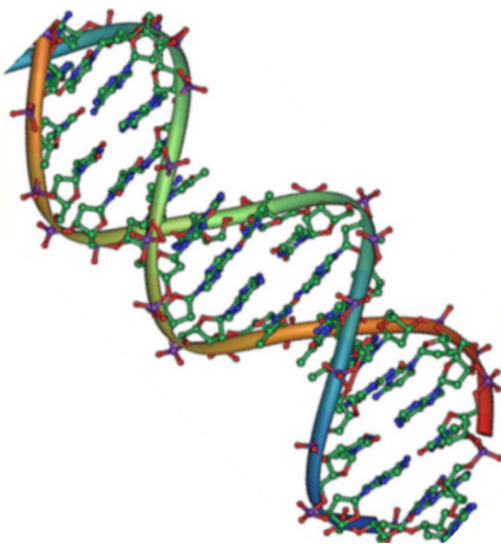
Adenine and guanine are classified as **purines**. The primary structure of a purine is two carbon-nitrogen rings. Cytosine, thymine, and uracil are classified as **pyrimidines** which have a single carbon-nitrogen ring as their primary structure ([\[link\]](#)). Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, the nitrogenous bases are simply known by their symbols A, T, G, C, and U. DNA contains A, T, G, and C whereas RNA contains A, U, G, and C.

The pentose sugar in DNA is deoxyribose, and in RNA, the sugar is ribose ([\[link\]](#)). The difference between the sugars is the presence of the hydroxyl group on the second carbon of the ribose and hydrogen on the second carbon of the deoxyribose. The carbon atoms of the sugar molecule are numbered as 1', 2', 3', 4', and 5' (1' is read as "one prime"). The phosphate residue is attached to the hydroxyl group of the 5' carbon of one sugar and the hydroxyl group of the 3' carbon of the sugar of the next nucleotide, which forms a 5'–3' **phosphodiester** linkage. The phosphodiester linkage is not formed by simple dehydration reaction like the other linkages

connecting monomers in macromolecules: its formation involves the removal of two phosphate groups. A polynucleotide may have thousands of such phosphodiester linkages.

## DNA Double-Helix Structure

DNA has a double-helix structure ([\[link\]](#)). The sugar and phosphate lie on the outside of the helix, forming the backbone of the DNA. The nitrogenous bases are stacked in the interior, like the steps of a staircase, in pairs; the pairs are bound to each other by hydrogen bonds. Every base pair in the double helix is separated from the next base pair by 0.34 nm. The two strands of the helix run in opposite directions, meaning that the 5' carbon end of one strand will face the 3' carbon end of its matching strand. (This is referred to as antiparallel orientation and is important to DNA replication and in many nucleic acid interactions.)



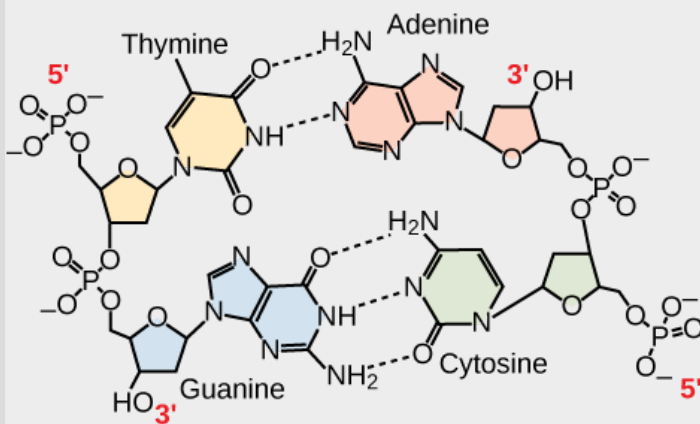
Native DNA is an antiparallel double helix. The phosphate backbone (indicated by the curvy lines) is on the outside, and the bases are on the

inside. Each base from one strand interacts via hydrogen bonding with a base from the opposing strand. (credit: Jerome Walker/Dennis Myts)

Only certain types of base pairing are allowed. For example, a certain purine can only pair with a certain pyrimidine. This means A can pair with T, and G can pair with C, as shown in [\[link\]](#). This is known as the base complementary rule. In other words, the DNA strands are complementary to each other. If the sequence of one strand is AATTGGCC, the complementary strand would have the sequence TTAACCGG. During DNA replication, each strand is copied, resulting in a daughter DNA double helix containing one parental DNA strand and a newly synthesized strand.

**Note:**

**Art Connection**



In a double stranded DNA molecule, the two strands run antiparallel to one another so that one strand runs 5' to 3' and the other 3' to 5'. The phosphate backbone is located on the outside,

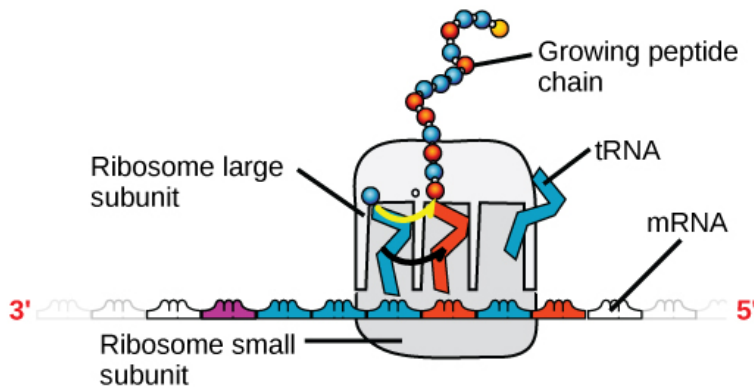
and the bases are in the middle.  
Adenine forms hydrogen bonds (or base pairs) with thymine, and guanine base pairs with cytosine.

A mutation occurs, and cytosine is replaced with adenine. What impact do you think this will have on the DNA structure?

## RNA

Ribonucleic acid, or RNA, is mainly involved in the process of protein synthesis under the direction of DNA. RNA is usually single-stranded and is made of ribonucleotides that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and the phosphate group.

There are four major types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and microRNA (miRNA). The first, mRNA, carries the message from DNA, which controls all of the cellular activities in a cell. If a cell requires a certain protein to be synthesized, the gene for this product is turned “on” and the messenger RNA is synthesized in the nucleus. The RNA base sequence is complementary to the coding sequence of the DNA from which it has been copied. However, in RNA, the base T is absent and U is present instead. If the DNA strand has a sequence AATTGCGC, the sequence of the complementary RNA is UUAACGCG. In the cytoplasm, the mRNA interacts with ribosomes and other cellular machinery ([link](#)).



A ribosome has two parts: a large subunit and a small subunit. The mRNA sits in between the two subunits. A tRNA molecule recognizes a codon on the mRNA, binds to it by complementary base pairing, and adds the correct amino acid to the growing peptide chain.

The mRNA is read in sets of three bases known as codons. Each codon codes for a single amino acid. In this way, the mRNA is read and the protein product is made. **Ribosomal RNA (rRNA)** is a major constituent of ribosomes on which the mRNA binds. The rRNA ensures the proper alignment of the mRNA and the ribosomes; the rRNA of the ribosome also has an enzymatic activity (peptidyl transferase) and catalyzes the formation of the peptide bonds between two aligned amino acids. **Transfer RNA (tRNA)** is one of the smallest of the four types of RNA, usually 70–90 nucleotides long. It carries the correct amino acid to the site of protein synthesis. It is the base pairing between the tRNA and mRNA that allows for the correct amino acid to be inserted in the polypeptide chain. microRNAs are the smallest RNA molecules and their role involves the regulation of gene expression by interfering with the expression of certain mRNA messages. [\[link\]](#) summarizes features of DNA and RNA.

Features of DNA and RNA		
	DNA	RNA
Function	Carries genetic information	Involved in protein synthesis
Location	Remains in the nucleus	Leaves the nucleus
Structure	Double helix	Usually single-stranded
Sugar	Deoxyribose	Ribose
Pyrimidines	Cytosine, thymine	Cytosine, uracil
Purines	Adenine, guanine	Adenine, guanine

Even though the RNA is single stranded, most RNA types show extensive intramolecular base pairing between complementary sequences, creating a predictable three-dimensional structure essential for their function.

As you have learned, information flow in an organism takes place from DNA to RNA to protein. DNA dictates the structure of mRNA in a process known as **transcription**, and RNA dictates the structure of protein in a process known as **translation**. This is known as the Central Dogma of Life, which holds true for all organisms; however, exceptions to the rule occur in connection with viral infections.

#### Note:

Link to Learning





To learn more about DNA, explore the [Howard Hughes Medical Institute BioInteractive animations](#) on the topic of DNA.

## Section Summary

Nucleic acids are molecules made up of nucleotides that direct cellular activities such as cell division and protein synthesis. Each nucleotide is made up of a pentose sugar, a nitrogenous base, and a phosphate group. There are two types of nucleic acids: DNA and RNA. DNA carries the genetic blueprint of the cell and is passed on from parents to offspring (in the form of chromosomes). It has a double-helical structure with the two strands running in opposite directions, connected by hydrogen bonds, and complementary to each other. RNA is single-stranded and is made of a pentose sugar (ribose), a nitrogenous base, and a phosphate group. RNA is involved in protein synthesis and its regulation. Messenger RNA (mRNA) is copied from the DNA, is exported from the nucleus to the cytoplasm, and contains information for the construction of proteins. Ribosomal RNA (rRNA) is a part of the ribosomes at the site of protein synthesis, whereas transfer RNA (tRNA) carries the amino acid to the site of protein synthesis. microRNA regulates the use of mRNA for protein synthesis.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) A mutation occurs, and cytosine is replaced with adenine. What impact do you think this will have on the DNA structure?

---

**Solution:**

[\[link\]](#) Adenine is larger than cytosine and will not be able to base pair properly with the guanine on the opposing strand. This will cause the DNA to bulge. DNA repair enzymes may recognize the bulge and replace the incorrect nucleotide.

**Review Questions****Exercise:**

**Problem:** A nucleotide of DNA may contain \_\_\_\_\_.

- a. ribose, uracil, and a phosphate group
- b. deoxyribose, uracil, and a phosphate group
- c. deoxyribose, thymine, and a phosphate group
- d. ribose, thymine, and a phosphate group

---

**Solution:**

C

**Exercise:**

**Problem:** The building blocks of nucleic acids are \_\_\_\_\_.

- a. sugars
- b. nitrogenous bases
- c. peptides
- d. nucleotides

---

**Solution:**

D

## Free Response

### Exercise:

**Problem:**What are the structural differences between RNA and DNA?

---

#### **Solution:**

DNA has a double-helix structure. The sugar and the phosphate are on the outside of the helix and the nitrogenous bases are in the interior. The monomers of DNA are nucleotides containing deoxyribose, one of the four nitrogenous bases (A, T, G and C), and a phosphate group. RNA is usually single-stranded and is made of ribonucleotides that are linked by phosphodiester linkages. A ribonucleotide contains ribose (the pentose sugar), one of the four nitrogenous bases (A,U, G, and C), and the phosphate group.

### Exercise:

**Problem:**What are the four types of RNA and how do they function?

---

#### **Solution:**

The four types of RNA are messenger RNA, ribosomal RNA, transfer RNA, and microRNA. Messenger RNA carries the information from the DNA that controls all cellular activities. The mRNA binds to the ribosomes that are constructed of proteins and rRNA, and tRNA transfers the correct amino acid to the site of protein synthesis. microRNA regulates the availability of mRNA for translation.

## Glossary

deoxyribonucleic acid (DNA)

double-helical molecule that carries the hereditary information of the cell

messenger RNA (mRNA)

RNA that carries information from DNA to ribosomes during protein synthesis

nucleic acid

biological macromolecule that carries the genetic blueprint of a cell and carries instructions for the functioning of the cell

nucleotide

monomer of nucleic acids; contains a pentose sugar, one or more phosphate groups, and a nitrogenous base

phosphodiester

linkage covalent chemical bond that holds together the polynucleotide chains with a phosphate group linking two pentose sugars of neighboring nucleotides

polynucleotide

long chain of nucleotides

purine

type of nitrogenous base in DNA and RNA; adenine and guanine are purines

pyrimidine

type of nitrogenous base in DNA and RNA; cytosine, thymine, and uracil are pyrimidines

ribonucleic acid (RNA)

single-stranded, often internally base paired, molecule that is involved in protein synthesis

ribosomal RNA (rRNA)

RNA that ensures the proper alignment of the mRNA and the ribosomes during protein synthesis and catalyzes the formation of the peptide linkage

transcription

process through which messenger RNA forms on a template of DNA

transfer RNA (tRNA)

RNA that carries activated amino acids to the site of protein synthesis on the ribosome

translation

process through which RNA directs the formation of protein

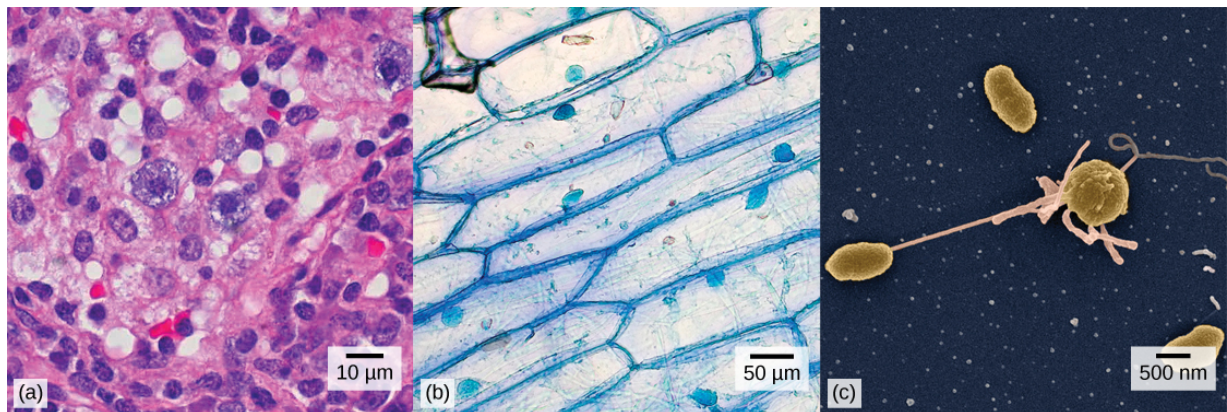
## Introduction

class="introduction"

(a) Nasal  
sinus cells  
(viewed with  
a light  
microscope),  
(b) onion  
cells (viewed  
with a light  
microscope),  
and (c) *Vibrio  
tasmaniensis*  
bacterial cells  
(seen through  
a scanning  
electron  
microscope)  
are from very  
different  
organisms,  
yet all share  
certain  
characteristic  
s of basic cell  
structure.

(credit a:  
modification  
of work by  
Ed Uthman,  
MD; credit b:  
modification  
of work by  
Umberto  
Salvagnin;  
credit c:

modification  
of work by  
Anthony  
D'Onofrio,  
William H.  
Fowle, Eric J.  
Stewart, and  
Kim Lewis of  
the Lewis  
Lab at  
Northeastern  
University;  
scale-bar data  
from Matt  
Russell)



Close your eyes and picture a brick wall. What is the basic building block of that wall? A single brick, of course. Like a brick wall, your body is composed of basic building blocks, and the building blocks of your body are cells.

Your body has many kinds of cells, each specialized for a specific purpose. Just as a home is made from a variety of building materials, the human body is constructed from many cell types. For example, epithelial cells protect the surface of the body and cover the organs and body cavities within. Bone cells help to support and protect the body. Cells of the immune

system fight invading bacteria. Additionally, blood and blood cells carry nutrients and oxygen throughout the body while removing carbon dioxide. Each of these cell types plays a vital role during the growth, development, and day-to-day maintenance of the body. In spite of their enormous variety, however, cells from all organisms—even ones as diverse as bacteria, onion, and human—share certain fundamental characteristics.



## Studying Cells

By the end of this section, you will be able to:

- Describe the role of cells in organisms
- Compare and contrast light microscopy and electron microscopy
- Summarize cell theory

A cell is the smallest unit of a living thing. A living thing, whether made of one cell (like bacteria) or many cells (like a human), is called an organism. Thus, cells are the basic building blocks of all organisms.

Several cells of one kind that interconnect with each other and perform a shared function form tissues, several tissues combine to form an organ (your stomach, heart, or brain), and several organs make up an organ system (such as the digestive system, circulatory system, or nervous system). Several systems that function together form an organism (like a human being). Here, we will examine the structure and function of cells.

There are many types of cells, all grouped into one of two broad categories: prokaryotic and eukaryotic. For example, both animal and plant cells are classified as eukaryotic cells, whereas bacterial cells are classified as prokaryotic. Before discussing the criteria for determining whether a cell is prokaryotic or eukaryotic, let's first examine how biologists study cells.

## Microscopy

Cells vary in size. With few exceptions, individual cells cannot be seen with the naked eye, so scientists use microscopes (micro- = “small”; -scope = “to look at”) to study them. A **microscope** is an instrument that magnifies an object. Most photographs of cells are taken with a microscope, and these images can also be called micrographs.

The optics of a microscope's lenses change the orientation of the image that the user sees. A specimen that is right-side up and facing right on the microscope slide will appear upside-down and facing left when viewed through a microscope, and vice versa. Similarly, if the slide is moved left while looking through the microscope, it will appear to move right, and if

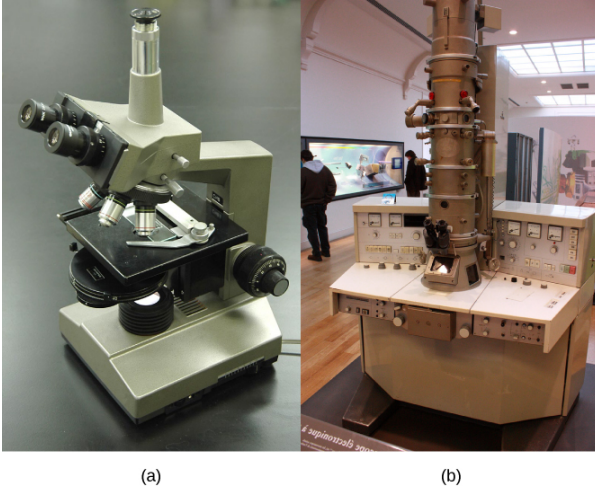
moved down, it will seem to move up. This occurs because microscopes use two sets of lenses to magnify the image. Because of the manner by which light travels through the lenses, this system of two lenses produces an inverted image (binocular, or dissecting microscopes, work in a similar manner, but include an additional magnification system that makes the final image appear to be upright).

## Light Microscopes

To give you a sense of cell size, a typical human red blood cell is about eight millionths of a meter or eight micrometers (abbreviated as eight  $\mu\text{m}$ ) in diameter; the head of a pin is about two thousandths of a meter (two mm) in diameter. That means about 250 red blood cells could fit on the head of a pin.

Most student microscopes are classified as **light microscopes** ([link](#)a). Visible light passes and is bent through the lens system to enable the user to see the specimen. Light microscopes are advantageous for viewing living organisms, but since individual cells are generally transparent, their components are not distinguishable unless they are colored with special stains. Staining, however, usually kills the cells.

Light microscopes commonly used in the undergraduate college laboratory magnify up to approximately 400 times. Two parameters that are important in microscopy are magnification and resolving power. Magnification is the process of enlarging an object in appearance. Resolving power is the ability of a microscope to distinguish two adjacent structures as separate: the higher the resolution, the better the clarity and detail of the image. When oil immersion lenses are used for the study of small objects, magnification is usually increased to 1,000 times. In order to gain a better understanding of cellular structure and function, scientists typically use electron microscopes.



(a) Most light microscopes used in a college biology lab can magnify cells up to approximately 400 times and have a resolution of about 200 nanometers. (b) Electron microscopes provide a much higher magnification, 100,000x, and have a resolution of 50 picometers. (credit a: modification of work by "GcG"/Wikimedia Commons; credit b: modification of work by Evan Bench)

## Electron Microscopes

In contrast to light microscopes, **electron microscopes** ([link](#)) use a beam of electrons instead of a beam of light. Not only does this allow for higher magnification and, thus, more detail ([link](#)), it also provides higher resolving power. The method used to prepare the specimen for viewing with an electron microscope kills the specimen. Electrons have short

wavelengths (shorter than photons) that move best in a vacuum, so living cells cannot be viewed with an electron microscope.

In a scanning electron microscope, a beam of electrons moves back and forth across a cell's surface, creating details of cell surface characteristics. In a transmission electron microscope, the electron beam penetrates the cell and provides details of a cell's internal structures. As you might imagine, electron microscopes are significantly more bulky and expensive than light microscopes.



(a) These *Salmonella* bacteria appear as tiny purple dots when viewed with a light microscope. (b) This scanning electron microscope micrograph shows *Salmonella* bacteria (in red) invading human cells (yellow). Even though subfigure (b) shows a different *Salmonella* specimen than subfigure (a), you can still observe the comparative increase in magnification and detail. (credit a: modification of work by CDC/Armed Forces Institute of Pathology, Charles N. Farmer, Rocky Mountain Laboratories; credit b: modification of work by NIAID, NIH; scale-bar data from Matt Russell)

**Note:**

Link to Learning



For another perspective on cell size, try the HowBig interactive at [this site](#).

## Cell Theory

The microscopes we use today are far more complex than those used in the 1600s by Antony van Leeuwenhoek, a Dutch shopkeeper who had great skill in crafting lenses. Despite the limitations of his now-ancient lenses, van Leeuwenhoek observed the movements of protista (a type of single-celled organism) and sperm, which he collectively termed “animalcules.”

In a 1665 publication called *Micrographia*, experimental scientist Robert Hooke coined the term “cell” for the box-like structures he observed when viewing cork tissue through a lens. In the 1670s, van Leeuwenhoek discovered bacteria and protozoa. Later advances in lenses, microscope construction, and staining techniques enabled other scientists to see some components inside cells.

By the late 1830s, botanist Matthias Schleiden and zoologist Theodor Schwann were studying tissues and proposed the **unified cell theory**, which states that all living things are composed of one or more cells, the cell is the basic unit of life, and new cells arise from existing cells. Rudolf Virchow later made important contributions to this theory.

**Note:**

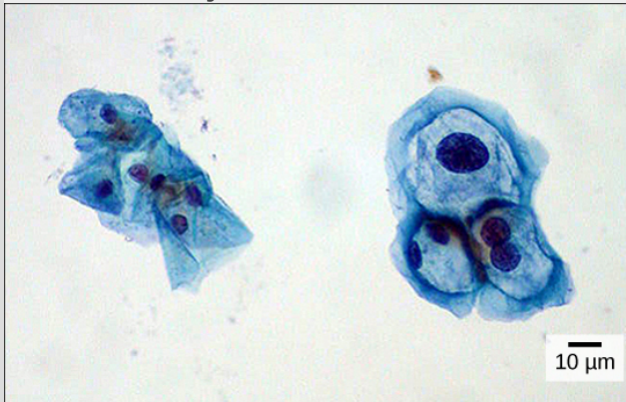
## Career Connection

### Cytotechnologist

Have you ever heard of a medical test called a Pap smear ([link](#))? In this test, a doctor takes a small sample of cells from the uterine cervix of a patient and sends it to a medical lab where a cytotechnologist stains the cells and examines them for any changes that could indicate cervical cancer or a microbial infection.

Cytotechnologists (cyto- = “cell”) are professionals who study cells via microscopic examinations and other laboratory tests. They are trained to determine which cellular changes are within normal limits and which are abnormal. Their focus is not limited to cervical cells; they study cellular specimens that come from all organs. When they notice abnormalities, they consult a pathologist, who is a medical doctor who can make a clinical diagnosis.

Cytotechnologists play a vital role in saving people’s lives. When abnormalities are discovered early, a patient’s treatment can begin sooner, which usually increases the chances of a successful outcome.



These uterine cervix cells, viewed through a light microscope, were obtained from a Pap smear.

Normal cells are on the left. The cells on the right are infected with human papillomavirus (HPV).

Notice that the infected cells are larger; also, two of these cells each have two nuclei instead of one, the normal number. (credit:

modification of work by Ed  
Uthman, MD; scale-bar data from  
Matt Russell)

## Section Summary

A cell is the smallest unit of life. Most cells are so tiny that they cannot be seen with the naked eye. Therefore, scientists use microscopes to study cells. Electron microscopes provide higher magnification, higher resolution, and more detail than light microscopes. The unified cell theory states that all organisms are composed of one or more cells, the cell is the basic unit of life, and new cells arise from existing cells.

## Review Questions

### Exercise:

#### Problem:

When viewing a specimen through a light microscope, scientists use \_\_\_\_\_ to distinguish the individual components of cells.

- a. a beam of electrons
- b. radioactive isotopes
- c. special stains
- d. high temperatures

---

#### Solution:

C

### Exercise:

**Problem:** The \_\_\_\_\_ is the basic unit of life.

- a. organism
- b. cell
- c. tissue
- d. organ

---

**Solution:**

B

## Free Response

**Exercise:**

**Problem:**

In your everyday life, you have probably noticed that certain instruments are ideal for certain situations. For example, you would use a spoon rather than a fork to eat soup because a spoon is shaped for scooping, while soup would slip between the tines of a fork. The use of ideal instruments also applies in science. In what situation(s) would the use of a light microscope be ideal, and why?

---

**Solution:**

A light microscope would be ideal when viewing a small living organism, especially when the cell has been stained to reveal details.

**Exercise:**

**Problem:**

In what situation(s) would the use of a scanning electron microscope be ideal, and why?



---

**Solution:**

A scanning electron microscope would be ideal when you want to view the minute details of a cell's surface, because its beam of electrons moves back and forth over the surface to convey the image.

**Exercise:****Problem:**

In what situation(s) would a transmission electron microscope be ideal, and why?

---

**Solution:**

A transmission electron microscope would be ideal for viewing the cell's internal structures, because many of the internal structures have membranes that are not visible by the light microscope.

**Exercise:****Problem:**

What are the advantages and disadvantages of each of these types of microscopes?

---

**Solution:**

The advantages of light microscopes are that they are easily obtained, and the light beam does not kill the cells. However, typical light microscopes are somewhat limited in the amount of detail they can reveal. Electron microscopes are ideal because you can view intricate details, but they are bulky and costly, and preparation for the microscopic examination kills the specimen.

**Glossary**

cell theory  
    see unified cell theory

electron microscope

an instrument that magnifies an object using a beam of electrons passed and bent through a lens system to visualize a specimen

light microscope

an instrument that magnifies an object using a beam visible light passed and bent through a lens system to visualize a specimen

microscope

an instrument that magnifies an object

unified cell theory

a biological concept that states that all organisms are composed of one or more cells; the cell is the basic unit of life; and new cells arise from existing cells

## Prokaryotic Cells

By the end of this section, you will be able to:

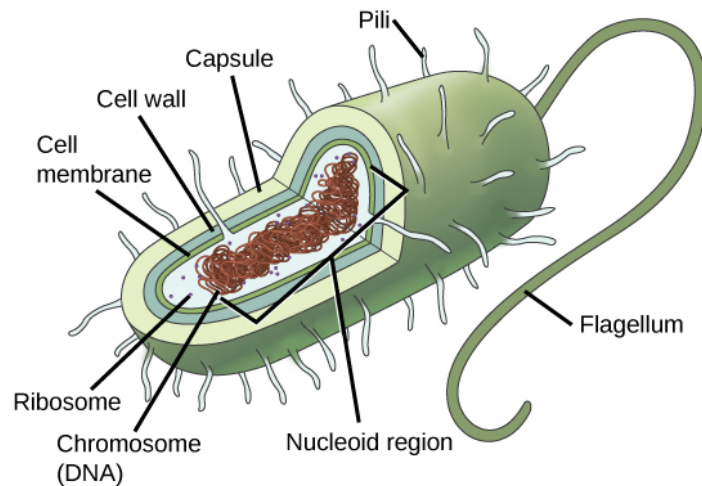
- Name examples of prokaryotic and eukaryotic organisms
- Compare and contrast prokaryotic cells and eukaryotic cells
- Describe the relative sizes of different kinds of cells
- Explain why cells must be small

Cells fall into one of two broad categories: prokaryotic and eukaryotic. Only the predominantly single-celled organisms of the domains Bacteria and Archaea are classified as prokaryotes (pro- = “before”; -kary- = “nucleus”). Animals, plants, fungi, and protists are all eukaryotes (eu- = “true”) and are made up of eukaryotic cells.

## Components of Prokaryotic Cells

All cells share four common components: 1) a plasma membrane, an outer covering that separates the cell’s interior from its surrounding environment; 2) cytoplasm, consisting of a jelly-like cytosol within the cell in which other cellular components are found; 3) DNA, the genetic material of the cell; and 4) ribosomes, which synthesize proteins. However, prokaryotes differ from eukaryotic cells in several ways.

A **prokaryote** is a simple, mostly single-celled (unicellular) organism that lacks a nucleus, or any other membrane-bound organelle. We will shortly come to see that this is significantly different in eukaryotes. Prokaryotic DNA is found in a central part of the cell: the **nucleoid** ([\[link\]](#)).



This figure shows the generalized structure of a prokaryotic cell. All prokaryotes have chromosomal DNA localized in a nucleoid, ribosomes, a cell membrane, and a cell wall. The other structures shown are present in some, but not all, bacteria.

Most prokaryotes have a peptidoglycan cell wall and many have a polysaccharide capsule ([link](#)). The cell wall acts as an extra layer of protection, helps the cell maintain its shape, and prevents dehydration. The capsule enables the cell to attach to surfaces in its environment. Some prokaryotes have flagella, pili, or fimbriae. Flagella are used for locomotion. Pili are used to exchange genetic material during a type of reproduction called conjugation. Fimbriae are used by bacteria to attach to a host cell.

**Note:****Career Connection****Microbiologist**

The most effective action anyone can take to prevent the spread of contagious illnesses is to wash his or her hands. Why? Because microbes

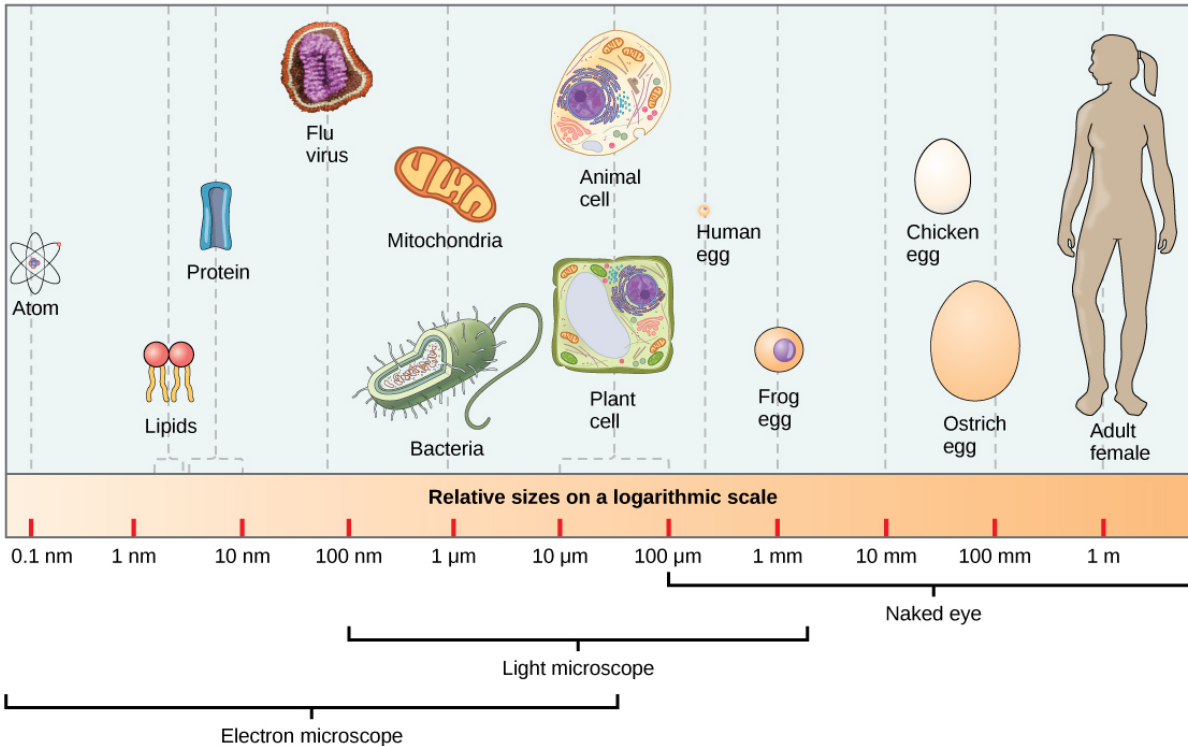
(organisms so tiny that they can only be seen with microscopes) are ubiquitous. They live on doorknobs, money, your hands, and many other surfaces. If someone sneezes into his hand and touches a doorknob, and afterwards you touch that same doorknob, the microbes from the sneezer's mucus are now on your hands. If you touch your hands to your mouth, nose, or eyes, those microbes can enter your body and could make you sick.

However, not all microbes (also called microorganisms) cause disease; most are actually beneficial. You have microbes in your gut that make vitamin K. Other microorganisms are used to ferment beer and wine. Microbiologists are scientists who study microbes. Microbiologists can pursue a number of careers. Not only do they work in the food industry, they are also employed in the veterinary and medical fields. They can work in the pharmaceutical sector, serving key roles in research and development by identifying new sources of antibiotics that could be used to treat bacterial infections.

Environmental microbiologists may look for new ways to use specially selected or genetically engineered microbes for the removal of pollutants from soil or groundwater, as well as hazardous elements from contaminated sites. These uses of microbes are called bioremediation technologies. Microbiologists can also work in the field of bioinformatics, providing specialized knowledge and insight for the design, development, and specificity of computer models of, for example, bacterial epidemics.

## Cell Size

At 0.1 to 5.0  $\mu\text{m}$  in diameter, prokaryotic cells are significantly smaller than eukaryotic cells, which have diameters ranging from 10 to 100  $\mu\text{m}$  ([link](#)). The small size of prokaryotes allows ions and organic molecules that enter them to quickly diffuse to other parts of the cell. Similarly, any wastes produced within a prokaryotic cell can quickly diffuse out. This is not the case in eukaryotic cells, which have developed different structural adaptations to enhance intracellular transport.



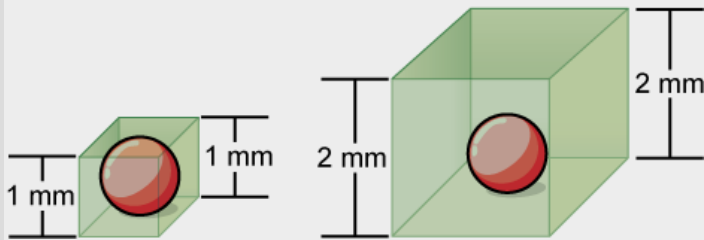
This figure shows relative sizes of microbes on a logarithmic scale (recall that each unit of increase in a logarithmic scale represents a 10-fold increase in the quantity being measured).

Small size, in general, is necessary for all cells, whether prokaryotic or eukaryotic. Let's examine why that is so. First, we'll consider the area and volume of a typical cell. Not all cells are spherical in shape, but most tend to approximate a sphere. You may remember from your high school geometry course that the formula for the surface area of a sphere is  $4\pi r^2$ , while the formula for its volume is  $\frac{4\pi r^3}{3}$ . Thus, as the radius of a cell increases, its surface area increases as the square of its radius, but its volume increases as the cube of its radius (much more rapidly). Therefore, as a cell increases in size, its surface area-to-volume ratio decreases. This same principle would apply if the cell had the shape of a cube ([\[link\]](#)). If the cell grows too large, the plasma membrane will not have sufficient surface area to support the rate of diffusion required for the increased volume. In other words, as a cell grows, it becomes less efficient. One way to become more efficient is to divide; another way is to develop organelles that

perform specific tasks. These adaptations lead to the development of more sophisticated cells called eukaryotic cells.

**Note:**

**Art Connection**



Notice that as a cell increases in size, its surface area-to-volume ratio decreases. When there is insufficient surface area to support a cell's increasing volume, a cell will either divide or die. The cell on the left has a volume of  $1 \text{ mm}^3$  and a surface area of  $6 \text{ mm}^2$ , with a surface area-to-volume ratio of 6 to 1, whereas the cell on the right has a volume of  $8 \text{ mm}^3$  and a surface area of  $24 \text{ mm}^2$ , with a surface area-to-volume ratio of 3 to 1.

Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

## Section Summary

Prokaryotes are predominantly single-celled organisms of the domains Bacteria and Archaea. All prokaryotes have plasma membranes, cytoplasm, ribosomes, and DNA that is not membrane-bound. Most have peptidoglycan cell walls and many have polysaccharide capsules. Prokaryotic cells range in diameter from 0.1 to 5.0  $\mu\text{m}$ .

As a cell increases in size, its surface area-to-volume ratio decreases. If the cell grows too large, the plasma membrane will not have sufficient surface area to support the rate of diffusion required for the increased volume.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

---

#### Solution:

[\[link\]](#) Substances can diffuse more quickly through small cells. Small cells have no need for organelles and therefore do not need to expend energy getting substances across organelle membranes. Large cells have organelles that can separate cellular processes, enabling them to build molecules that are more complex.

## Review Questions

### Exercise:

#### Problem:

Prokaryotes depend on \_\_\_\_\_ to obtain some materials and to get rid of wastes.

- a. ribosomes



- b. flagella
- c. cell division
- d. diffusion

---

**Solution:**

D

**Exercise:**

**Problem:** Bacteria that lack fimbriae are less likely to \_\_\_\_\_.

- a. adhere to cell surfaces
- b. swim through bodily fluids
- c. synthesize proteins
- d. retain the ability to divide

---

**Solution:**

A

**Free Response**

**Exercise:**

**Problem:**

Antibiotics are medicines that are used to fight bacterial infections. These medicines kill prokaryotic cells without harming human cells. What part or parts of the bacterial cell do you think antibiotics target? Why?

---

**Solution:**

The cell wall would be targeted by antibiotics as well as the bacteria's ability to replicate. This would inhibit the bacteria's ability to

reproduce, and it would compromise its defense mechanisms.

**Exercise:**

**Problem:** Explain why not all microbes are harmful.

---

**Solution:**

Some microbes are beneficial. For instance, *E. coli* bacteria populate the human gut and help break down fiber in the diet. Some foods such as yogurt are formed by bacteria.

## Glossary

nucleoid

central part of a prokaryotic cell in which the chromosome is found

prokaryote

unicellular organism that lacks a nucleus or any other membrane-bound organelle

## Eukaryotic Cells

By the end of this section, you will be able to:

- Describe the structure of eukaryotic cells
- Compare animal cells with plant cells
- State the role of the plasma membrane
- Summarize the functions of the major cell organelles

Have you ever heard the phrase “form follows function?” It’s a philosophy practiced in many industries. In architecture, this means that buildings should be constructed to support the activities that will be carried out inside them. For example, a skyscraper should be built with several elevator banks; a hospital should be built so that its emergency room is easily accessible.

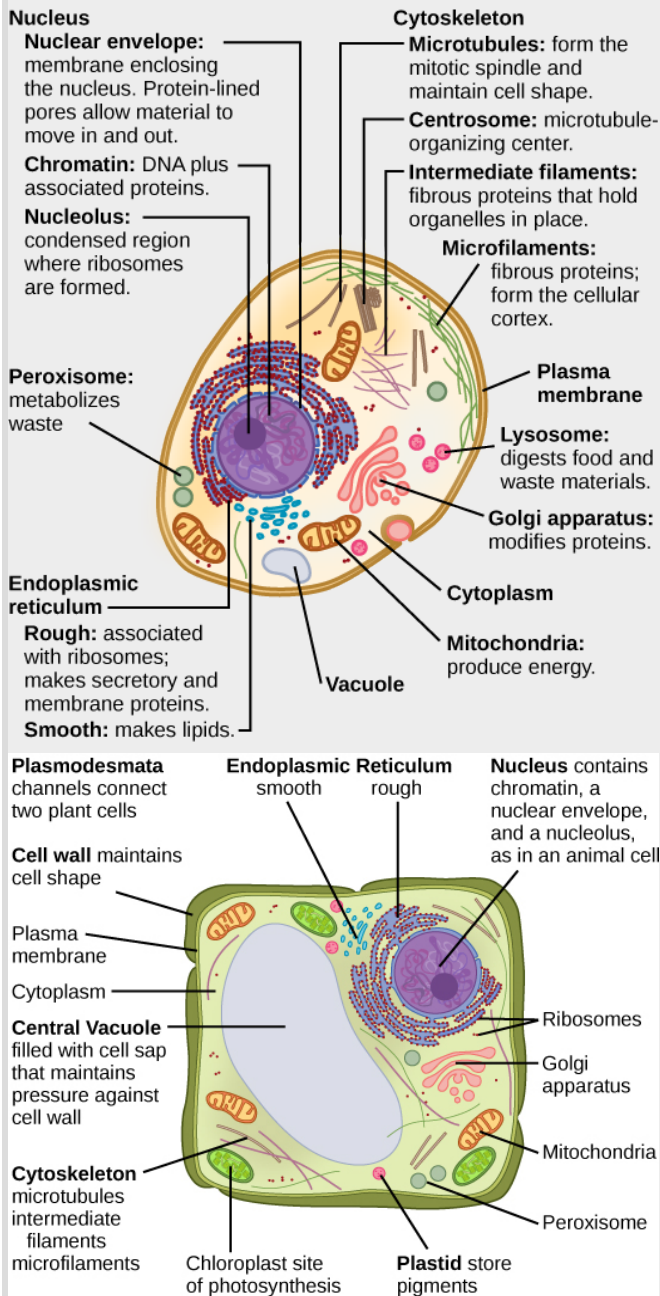
Our natural world also utilizes the principle of form following function, especially in cell biology, and this will become clear as we explore eukaryotic cells ( [link](#)). Unlike prokaryotic cells, **eukaryotic cells** have: 1) a membrane-bound nucleus; 2) numerous membrane-bound **organelles** such as the endoplasmic reticulum, Golgi apparatus, chloroplasts, mitochondria, and others; and 3) several, rod-shaped chromosomes. Because a eukaryotic cell’s nucleus is surrounded by a membrane, it is often said to have a “true nucleus.” The word “organelle” means “little organ,” and, as already mentioned, organelles have specialized cellular functions, just as the organs of your body have specialized functions.

At this point, it should be clear to you that eukaryotic cells have a more complex structure than prokaryotic cells. Organelles allow different functions to be compartmentalized in different areas of the cell. Before turning to organelles, let’s first examine two important components of the cell: the plasma membrane and the cytoplasm.

### **Note:**

Art Connection

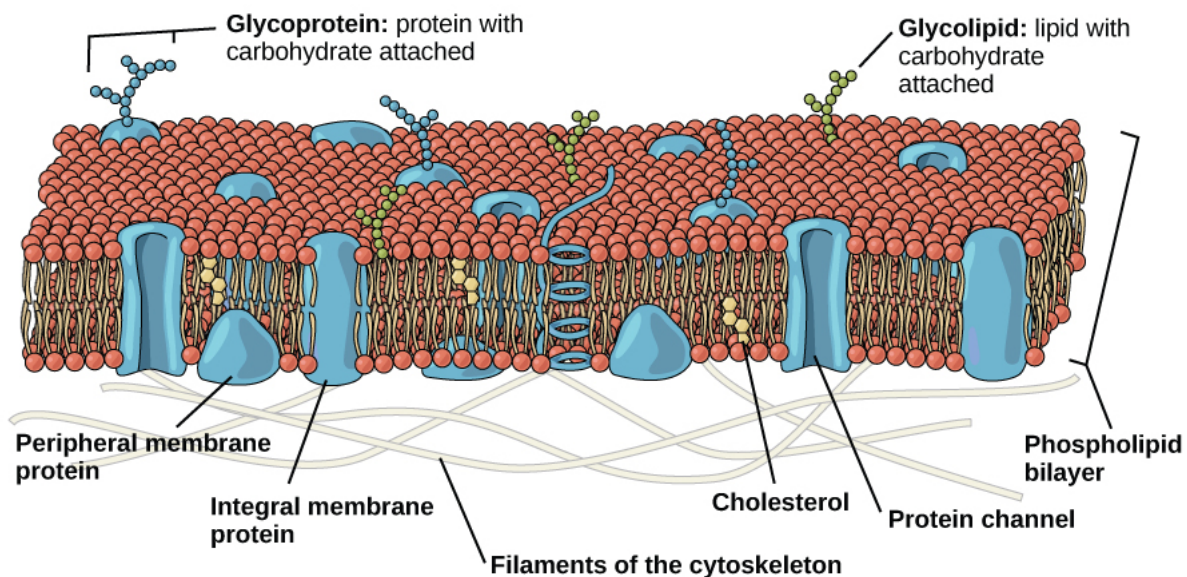
These figures show the major organelles and other cell components of (a) a typical animal cell and (b) a typical eukaryotic plant cell. The plant cell has a cell wall, chloroplasts, plastids, and a central vacuole —structures not found in animal cells. Plant cells do not have lysosomes or centrosomes.



If the nucleolus were not able to carry out its function, what other cellular organelles would be affected?

## The Plasma Membrane

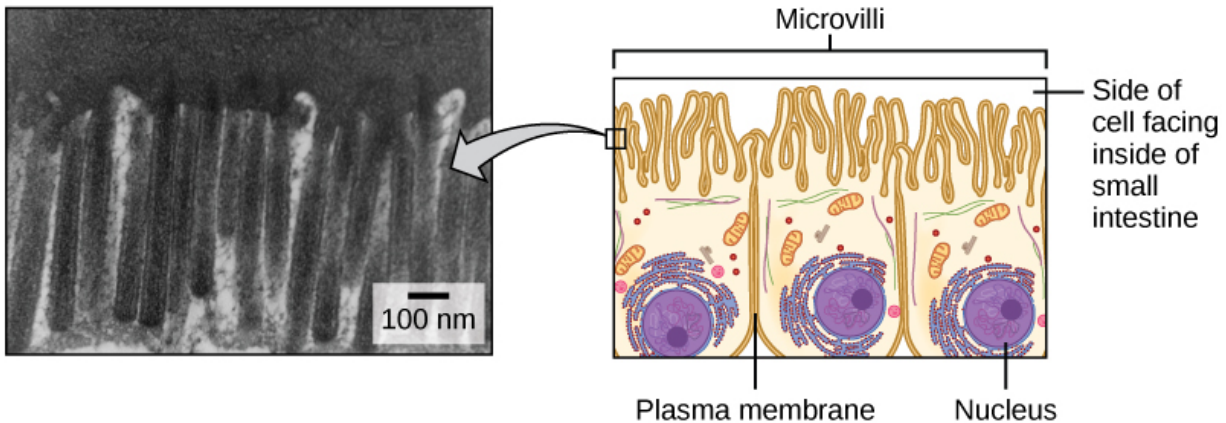
Like prokaryotes, eukaryotic cells have a **plasma membrane** ( [\[link\]](#)), a phospholipid bilayer with embedded proteins that separates the internal contents of the cell from its surrounding environment. A phospholipid is a lipid molecule with two fatty acid chains and a phosphate-containing group. The plasma membrane controls the passage of organic molecules, ions, water, and oxygen into and out of the cell. Wastes (such as carbon dioxide and ammonia) also leave the cell by passing through the plasma membrane.



The eukaryotic plasma membrane is a phospholipid bilayer with proteins and cholesterol embedded in it.

The plasma membranes of cells that specialize in absorption are folded into fingerlike projections called microvilli (singular = microvillus); ( [\[link\]](#)). Such cells are typically found lining the small intestine, the organ that absorbs nutrients from digested food. This is an excellent example of form following function. People with celiac disease have an immune response to gluten, which is a protein found in wheat, barley, and rye. The immune response damages microvilli, and thus, afflicted individuals cannot absorb

nutrients. This leads to malnutrition, cramping, and diarrhea. Patients suffering from celiac disease must follow a gluten-free diet.



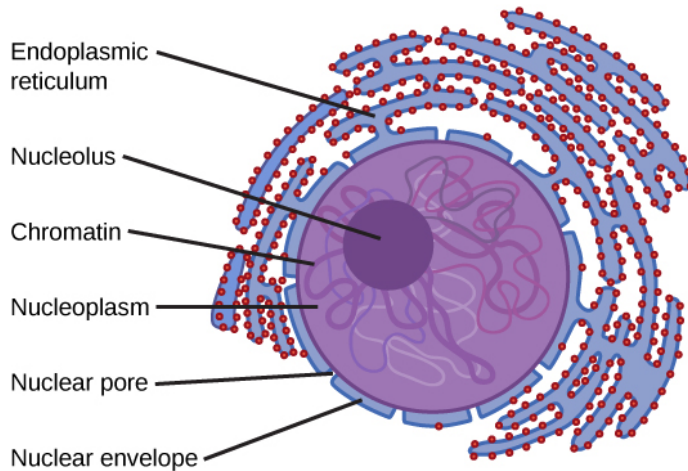
Microvilli, shown here as they appear on cells lining the small intestine, increase the surface area available for absorption. These microvilli are only found on the area of the plasma membrane that faces the cavity from which substances will be absorbed. (credit "micrograph": modification of work by Louisa Howard)

## The Cytoplasm

The **cytoplasm** is the entire region of a cell between the plasma membrane and the nuclear envelope (a structure to be discussed shortly). It is made up of organelles suspended in the gel-like **cytosol**, the cytoskeleton, and various chemicals ( [\[link\]](#)). Even though the cytoplasm consists of 70 to 80 percent water, it has a semi-solid consistency, which comes from the proteins within it. However, proteins are not the only organic molecules found in the cytoplasm. Glucose and other simple sugars, polysaccharides, amino acids, nucleic acids, fatty acids, and derivatives of glycerol are found there, too. Ions of sodium, potassium, calcium, and many other elements are also dissolved in the cytoplasm. Many metabolic reactions, including protein synthesis, take place in the cytoplasm.

## The Nucleus

Typically, the nucleus is the most prominent organelle in a cell ( [\[link\]](#)). The **nucleus** (plural = nuclei) houses the cell's DNA and directs the synthesis of ribosomes and proteins. Let's look at it in more detail ( [\[link\]](#)).



The nucleus stores chromatin (DNA plus proteins) in a gel-like substance called the nucleoplasm. The nucleolus is a condensed region of chromatin where ribosome synthesis occurs. The boundary of the nucleus is called the nuclear envelope. It consists of two phospholipid bilayers: an outer membrane and an inner membrane. The nuclear membrane is continuous with the endoplasmic reticulum. Nuclear pores allow substances to enter and exit the nucleus.

## The Nuclear Envelope

The **nuclear envelope** is a double-membrane structure that constitutes the outermost portion of the nucleus ( [\[link\]](#)). Both the inner and outer membranes of the nuclear envelope are phospholipid bilayers.

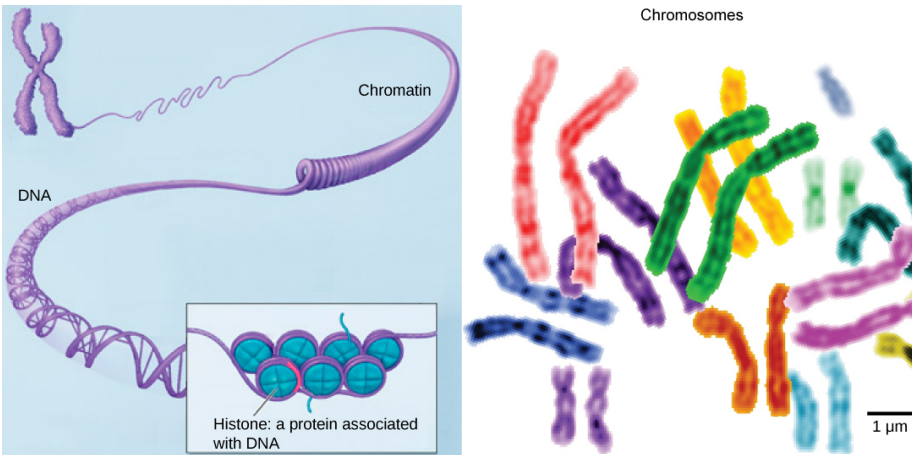
The nuclear envelope is punctuated with pores that control the passage of ions, molecules, and RNA between the nucleoplasm and cytoplasm. The **nucleoplasm** is the semi-solid fluid inside the nucleus, where we find the chromatin and the nucleolus.

## Chromatin and Chromosomes

To understand chromatin, it is helpful to first consider chromosomes.

**Chromosomes** are structures within the nucleus that are made up of DNA, the hereditary material. You may remember that in prokaryotes, DNA is organized into a single circular chromosome. In eukaryotes, chromosomes are linear structures. Every eukaryotic species has a specific number of chromosomes in the nuclei of its body's cells. For example, in humans, the chromosome number is 46, while in fruit flies, it is eight. Chromosomes are only visible and distinguishable from one another when the cell is getting ready to divide. When the cell is in the growth and maintenance phases of its life cycle, proteins are attached to chromosomes, and they resemble an unwound, jumbled bunch of threads. These unwound protein-chromosome complexes are called **chromatin** ( [\[link\]](#)); chromatin describes the material that makes up the chromosomes both when condensed and decondensed.





(a) This image shows various levels of the organization of chromatin (DNA and protein). (b) This image shows paired chromosomes. (credit b: modification of work by NIH; scale-bar data from Matt Russell)

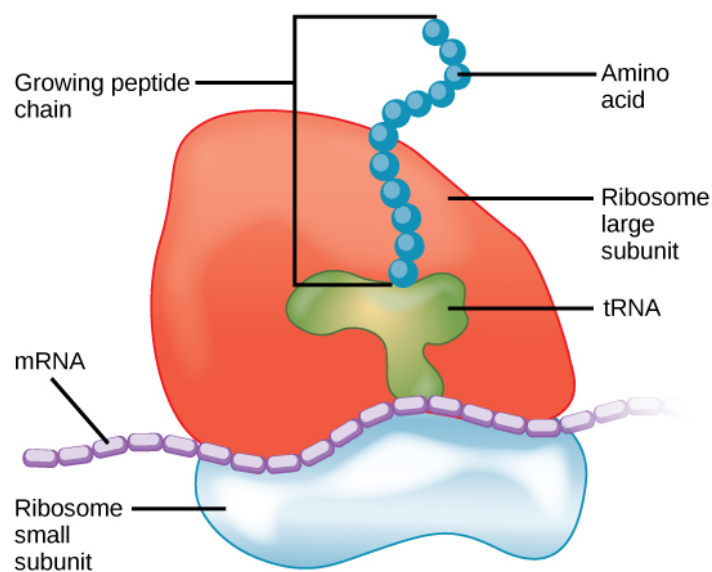
## The Nucleolus

We already know that the nucleus directs the synthesis of ribosomes, but how does it do this? Some chromosomes have sections of DNA that encode ribosomal RNA. A darkly staining area within the nucleus called the **nucleolus** (plural = nucleoli) aggregates the ribosomal RNA with associated proteins to assemble the ribosomal subunits that are then transported out through the pores in the nuclear envelope to the cytoplasm.

## Ribosomes

**Ribosomes** are the cellular structures responsible for protein synthesis. When viewed through an electron microscope, ribosomes appear either as clusters (polyribosomes) or single, tiny dots that float freely in the cytoplasm. They may be attached to the cytoplasmic side of the plasma membrane or the cytoplasmic side of the endoplasmic reticulum and the

outer membrane of the nuclear envelope ( [\[link\]](#)). Electron microscopy has shown us that ribosomes, which are large complexes of protein and RNA, consist of two subunits, aptly called large and small ( [\[link\]](#)). Ribosomes receive their “orders” for protein synthesis from the nucleus where the DNA is transcribed into messenger RNA (mRNA). The mRNA travels to the ribosomes, which translate the code provided by the sequence of the nitrogenous bases in the mRNA into a specific order of amino acids in a protein. Amino acids are the building blocks of proteins.



Ribosomes are made up of a large subunit (top) and a small subunit (bottom). During protein synthesis, ribosomes assemble amino acids into proteins.

Because proteins synthesis is an essential function of all cells (including enzymes, hormones, antibodies, pigments, structural components, and surface receptors), ribosomes are found in practically every cell. Ribosomes are particularly abundant in cells that synthesize large amounts of protein. For example, the pancreas is responsible for creating several digestive

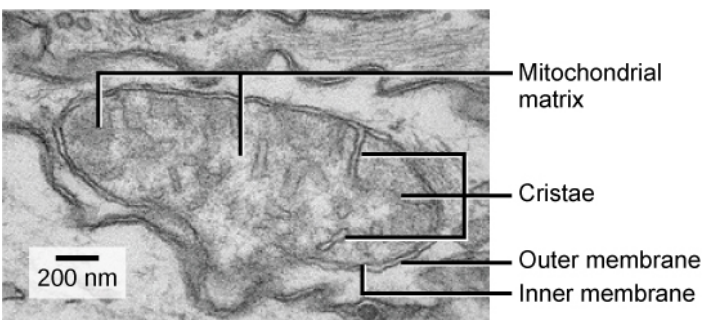
enzymes and the cells that produce these enzymes contain many ribosomes. Thus, we see another example of form following function.

## Mitochondria

**Mitochondria** (singular = mitochondrion) are often called the “powerhouses” or “energy factories” of a cell because they are responsible for making adenosine triphosphate (ATP), the cell’s main energy-carrying molecule. ATP represents the short-term stored energy of the cell. Cellular respiration is the process of making ATP using the chemical energy found in glucose and other nutrients. In mitochondria, this process uses oxygen and produces carbon dioxide as a waste product. In fact, the carbon dioxide that you exhale with every breath comes from the cellular reactions that produce carbon dioxide as a byproduct.

In keeping with our theme of form following function, it is important to point out that muscle cells have a very high concentration of mitochondria that produce ATP. Your muscle cells need a lot of energy to keep your body moving. When your cells don’t get enough oxygen, they do not make a lot of ATP. Instead, the small amount of ATP they make in the absence of oxygen is accompanied by the production of lactic acid.

Mitochondria are oval-shaped, double membrane organelles ( [\[link\]](#)) that have their own ribosomes and DNA. Each membrane is a phospholipid bilayer embedded with proteins. The inner layer has folds called cristae. The area surrounded by the folds is called the mitochondrial matrix. The cristae and the matrix have different roles in cellular respiration.



This electron micrograph shows a mitochondrion as viewed with a transmission electron microscope. This organelle has an outer membrane and an inner membrane. The inner membrane contains folds, called cristae, which increase its surface area. The space between the two membranes is called the intermembrane space, and the space inside the inner membrane is called the mitochondrial matrix. ATP synthesis takes place on the inner membrane. (credit: modification of work by Matthew Britton; scale-bar data from Matt Russell)

## **Peroxisomes**

**Peroxisomes** are small, round organelles enclosed by single membranes. They carry out oxidation reactions that break down fatty acids and amino acids. They also detoxify many poisons that may enter the body. (Many of these oxidation reactions release hydrogen peroxide,  $\text{H}_2\text{O}_2$ , which would be damaging to cells; however, when these reactions are confined to peroxisomes, enzymes safely break down the  $\text{H}_2\text{O}_2$  into oxygen and water.) For example, alcohol is detoxified by peroxisomes in liver cells. Glyoxysomes, which are specialized peroxisomes in plants, are responsible for converting stored fats into sugars.

## **Vesicles and Vacuoles**

**Vesicles** and **vacuoles** are membrane-bound sacs that function in storage and transport. Other than the fact that vacuoles are somewhat larger than vesicles, there is a very subtle distinction between them: The membranes of

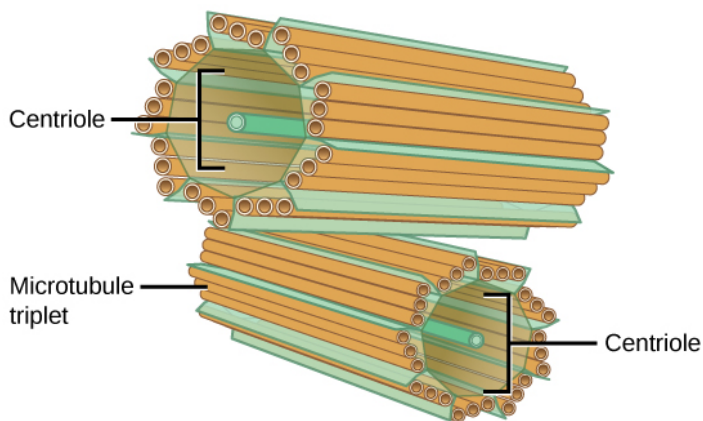
vesicles can fuse with either the plasma membrane or other membrane systems within the cell. Additionally, some agents such as enzymes within plant vacuoles break down macromolecules. The membrane of a vacuole does not fuse with the membranes of other cellular components.

## Animal Cells versus Plant Cells

At this point, you know that each eukaryotic cell has a plasma membrane, cytoplasm, a nucleus, ribosomes, mitochondria, peroxisomes, and in some, vacuoles, but there are some striking differences between animal and plant cells. While both animal and plant cells have microtubule organizing centers (MTOCs), animal cells also have centrioles associated with the MTOC: a complex called the centrosome. Animal cells each have a centrosome and lysosomes, whereas plant cells do not. Plant cells have a cell wall, chloroplasts and other specialized plastids, and a large central vacuole, whereas animal cells do not.

### The Centrosome

The **centrosome** is a microtubule-organizing center found near the nuclei of animal cells. It contains a pair of centrioles, two structures that lie perpendicular to each other ( [\[link\]](#)). Each centriole is a cylinder of nine triplets of microtubules.



The centrosome consists of two centrioles that lie at right angles to each other. Each centriole is a cylinder made up of nine triplets of microtubules. Nontubulin proteins (indicated by the green lines) hold the microtubule triplets together.

The centrosome (the organelle where all microtubules originate) replicates itself before a cell divides, and the centrioles appear to have some role in pulling the duplicated chromosomes to opposite ends of the dividing cell. However, the exact function of the centrioles in cell division isn't clear, because cells that have had the centrosome removed can still divide, and plant cells, which lack centrosomes, are capable of cell division.

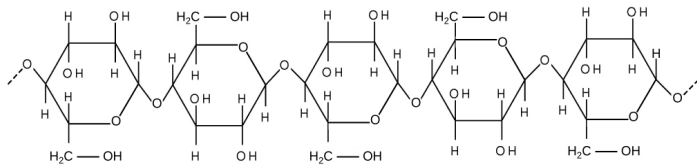
## **Lysosomes**

Animal cells have another set of organelles not found in plant cells: lysosomes. The **lysosomes** are the cell's "garbage disposal." In plant cells, the digestive processes take place in vacuoles. Enzymes within the lysosomes aid the breakdown of proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles. These enzymes are active at a much lower pH than that of the cytoplasm. Therefore, the pH within lysosomes is more acidic than the pH of the cytoplasm. Many reactions that take place in the cytoplasm could not occur at a low pH, so again, the advantage of compartmentalizing the eukaryotic cell into organelles is apparent.

## **The Cell Wall**

If you examine [\[link\]](#)**b**, the diagram of a plant cell, you will see a structure external to the plasma membrane called the cell wall. The **cell wall** is a rigid covering that protects the cell, provides structural support, and gives

shape to the cell. Fungal and protistan cells also have cell walls. While the chief component of prokaryotic cell walls is peptidoglycan, the major organic molecule in the plant cell wall is cellulose ( [link](#)), a polysaccharide made up of glucose units. Have you ever noticed that when you bite into a raw vegetable, like celery, it crunches? That's because you are tearing the rigid cell walls of the celery cells with your teeth.



Cellulose is a long chain of  $\beta$ -glucose molecules connected by a 1-4 linkage.

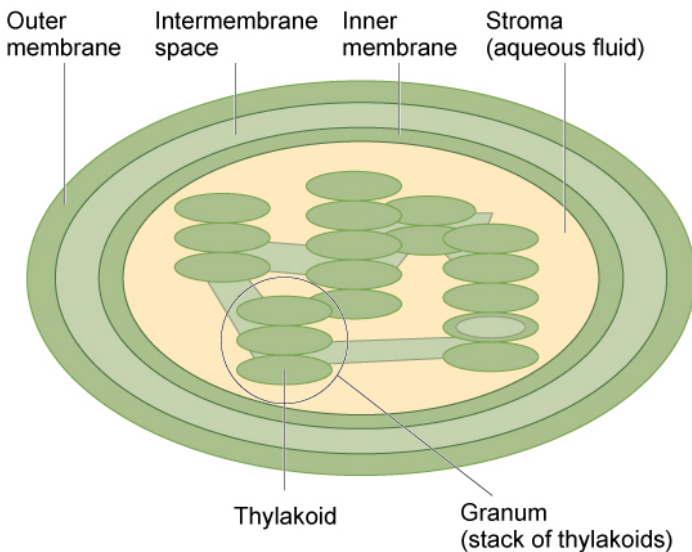
The dashed lines at each end of the figure indicate a series of many more glucose units. The size of the page makes it impossible to portray an entire cellulose molecule.

## Chloroplasts

Like the mitochondria, chloroplasts have their own DNA and ribosomes, but chloroplasts have an entirely different function. **Chloroplasts** are plant cell organelles that carry out photosynthesis. Photosynthesis is the series of reactions that use carbon dioxide, water, and light energy to make glucose and oxygen. This is a major difference between plants and animals; plants (autotrophs) are able to make their own food, like sugars, while animals (heterotrophs) must ingest their food.

Like mitochondria, chloroplasts have outer and inner membranes, but within the space enclosed by a chloroplast's inner membrane is a set of

interconnected and stacked fluid-filled membrane sacs called thylakoids ([\[link\]](#)). Each stack of thylakoids is called a granum (plural = grana). The fluid enclosed by the inner membrane that surrounds the grana is called the stroma.



The chloroplast has an outer membrane, an inner membrane, and membrane structures called thylakoids that are stacked into grana. The space inside the thylakoid membranes is called the thylakoid space. The light harvesting reactions take place in the thylakoid membranes, and the synthesis of sugar takes place in the fluid inside the inner membrane, which is called the stroma. Chloroplasts also have their own genome, which is contained on a single circular chromosome.



The chloroplasts contain a green pigment called **chlorophyll**, which captures the light energy that drives the reactions of photosynthesis. Like plant cells, photosynthetic protists also have chloroplasts. Some bacteria perform photosynthesis, but their chlorophyll is not relegated to an organelle.

**Note:**

**Evolution Connection**

**Endosymbiosis**

We have mentioned that both mitochondria and chloroplasts contain DNA and ribosomes. Have you wondered why? Strong evidence points to endosymbiosis as the explanation.

Symbiosis is a relationship in which organisms from two separate species depend on each other for their survival. Endosymbiosis (endo- = “within”) is a mutually beneficial relationship in which one organism lives inside the other. Endosymbiotic relationships abound in nature. We have already mentioned that microbes that produce vitamin K live inside the human gut. This relationship is beneficial for us because we are unable to synthesize vitamin K. It is also beneficial for the microbes because they are protected from other organisms and from drying out, and they receive abundant food from the environment of the large intestine.

Scientists have long noticed that bacteria, mitochondria, and chloroplasts are similar in size. We also know that bacteria have DNA and ribosomes, just as mitochondria and chloroplasts do. Scientists believe that host cells and bacteria formed an endosymbiotic relationship when the host cells ingested both aerobic and autotrophic bacteria (cyanobacteria) but did not destroy them. Through many millions of years of evolution, these ingested bacteria became more specialized in their functions, with the aerobic bacteria becoming mitochondria and the autotrophic bacteria becoming chloroplasts.

**The Central Vacuole**

Previously, we mentioned vacuoles as essential components of plant cells. If you look at [link](#)b, you will see that plant cells each have a large central vacuole that occupies most of the area of the cell. The **central vacuole** plays a key role in regulating the cell's concentration of water in changing environmental conditions. Have you ever noticed that if you forget to water a plant for a few days, it wilts? That's because as the water concentration in the soil becomes lower than the water concentration in the plant, water moves out of the central vacuoles and cytoplasm. As the central vacuole shrinks, it leaves the cell wall unsupported. This loss of support to the cell walls of plant cells results in the wilted appearance of the plant.

The central vacuole also supports the expansion of the cell. When the central vacuole holds more water, the cell gets larger without having to invest a lot of energy in synthesizing new cytoplasm.

## Section Summary

Like a prokaryotic cell, a eukaryotic cell has a plasma membrane, cytoplasm, and ribosomes, but a eukaryotic cell is typically larger than a prokaryotic cell, has a true nucleus (meaning its DNA is surrounded by a membrane), and has other membrane-bound organelles that allow for compartmentalization of functions. The plasma membrane is a phospholipid bilayer embedded with proteins. The nucleus's nucleolus is the site of ribosome assembly. Ribosomes are either found in the cytoplasm or attached to the cytoplasmic side of the plasma membrane or endoplasmic reticulum. They perform protein synthesis. Mitochondria participate in cellular respiration; they are responsible for the majority of ATP produced in the cell. Peroxisomes hydrolyze fatty acids, amino acids, and some toxins. Vesicles and vacuoles are storage and transport compartments. In plant cells, vacuoles also help break down macromolecules.

Animal cells also have a centrosome and lysosomes. The centrosome has two bodies perpendicular to each other, the centrioles, and has an unknown purpose in cell division. Lysosomes are the digestive organelles of animal cells.

Plant cells and plant-like cells each have a cell wall, chloroplasts, and a central vacuole. The plant cell wall, whose primary component is cellulose, protects the cell, provides structural support, and gives shape to the cell. Photosynthesis takes place in chloroplasts. The central vacuole can expand without having to produce more cytoplasm.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) If the nucleolus were not able to carry out its function, what other cellular organelles would be affected?

---

#### Solution:

[\[link\]](#) Free ribosomes and rough endoplasmic reticulum (which contains ribosomes) would not be able to form.

## Review Questions

### Exercise:

#### Problem:

Which of the following is surrounded by two phospholipid bilayers?

- a. the ribosomes
- b. the vesicles
- c. the cytoplasm
- d. the nucleoplasm

---

#### Solution:

D

**Exercise:**

**Problem:** Peroxisomes got their name because hydrogen peroxide is:

- a. used in their detoxification reactions
- b. produced during their oxidation reactions
- c. incorporated into their membranes
- d. a cofactor for the organelles' enzymes

---

**Solution:**

B

**Exercise:**

**Problem:**

In plant cells, the function of the lysosomes is carried out by  
\_\_\_\_\_.

- a. vacuoles
- b. peroxisomes
- c. ribosomes
- d. nuclei

---

**Solution:**

A

**Exercise:**

**Problem:**

Which of the following is found both in eukaryotic and prokaryotic cells?

- a. nucleus
- b. mitochondrion

- c. vacuole
  - d. ribosomes
- 

**Solution:**

D

## Free Response

**Exercise:**

**Problem:**

You already know that ribosomes are abundant in red blood cells. In what other cells of the body would you find them in great abundance? Why?

---

**Solution:**

Ribosomes are abundant in muscle cells as well because muscle cells are constructed of the proteins made by the ribosomes.

**Exercise:**

**Problem:**

What are the structural and functional similarities and differences between mitochondria and chloroplasts?

---

**Solution:**

Both are similar in that they are enveloped in a double membrane, both have an intermembrane space, and both make ATP. Both mitochondria and chloroplasts have DNA, and mitochondria have inner folds called cristae and a matrix, while chloroplasts have chlorophyll and accessory pigments in the thylakoids that form stacks (grana) and a stroma.

## **Glossary**

### **cell wall**

rigid cell covering made of various molecules that protects the cell, provides structural support, and gives shape to the cell

### **central vacuole**

large plant cell organelle that regulates the cell's storage compartment, holds water, and plays a significant role in cell growth as the site of macromolecule degradation

### **centrosome**

region in animal cells made of two centrioles

### **chlorophyll**

green pigment that captures the light energy that drives the light reactions of photosynthesis

### **chloroplast**

plant cell organelle that carries out photosynthesis

### **chromatin**

protein-DNA complex that serves as the building material of chromosomes

### **chromosome**

structure within the nucleus that is made up of chromatin that contains DNA, the hereditary material

### **cytoplasm**

entire region between the plasma membrane and the nuclear envelope, consisting of organelles suspended in the gel-like cytosol, the cytoskeleton, and various chemicals

### **cytosol**

gel-like material of the cytoplasm in which cell structures are suspended

eukaryotic cell

cell that has a membrane-bound nucleus and several other membrane-bound compartments or sacs

lysosome

organelle in an animal cell that functions as the cell's digestive component; it breaks down proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles

mitochondria

(singular = mitochondrion) cellular organelles responsible for carrying out cellular respiration, resulting in the production of ATP, the cell's main energy-carrying molecule

nuclear envelope

double-membrane structure that constitutes the outermost portion of the nucleus

nucleolus

darkly staining body within the nucleus that is responsible for assembling the subunits of the ribosomes

nucleoplasm

semi-solid fluid inside the nucleus that contains the chromatin and nucleolus

nucleus

cell organelle that houses the cell's DNA and directs the synthesis of ribosomes and proteins

organelle

compartment or sac within a cell

peroxisome

small, round organelle that contains hydrogen peroxide, oxidizes fatty acids and amino acids, and detoxifies many poisons

plasma membrane

phospholipid bilayer with embedded (integral) or attached (peripheral) proteins, and separates the internal content of the cell from its surrounding environment

ribosome

cellular structure that carries out protein synthesis

vacuole

membrane-bound sac, somewhat larger than a vesicle, which functions in cellular storage and transport

vesicle

small, membrane-bound sac that functions in cellular storage and transport; its membrane is capable of fusing with the plasma membrane and the membranes of the endoplasmic reticulum and Golgi apparatus



## The Endomembrane System and Proteins

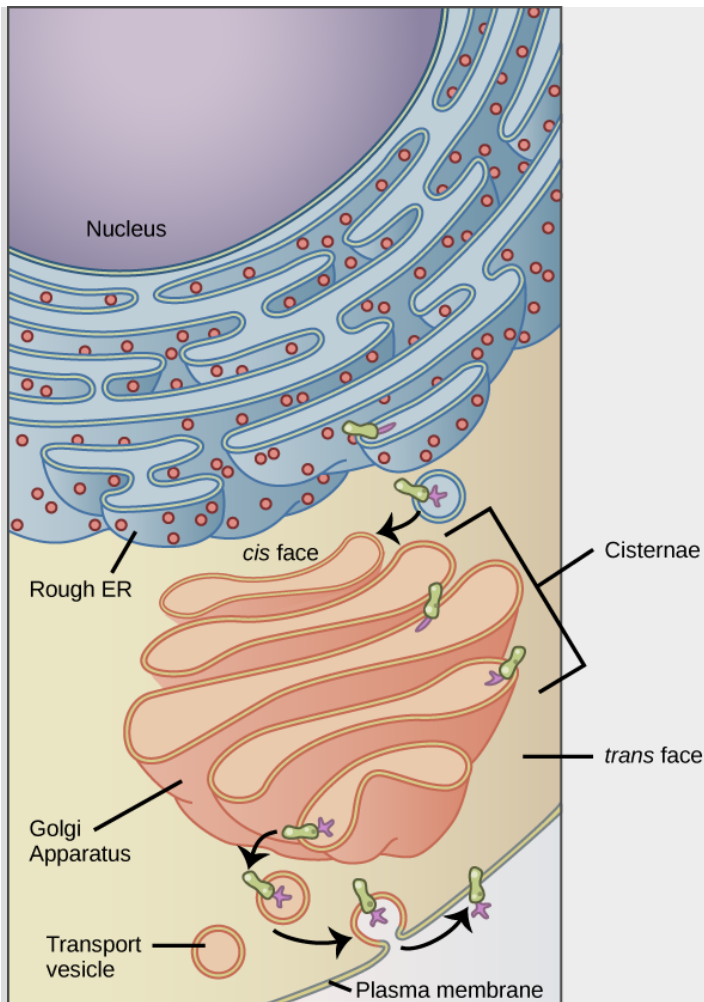
By the end of this section, you will be able to:

- List the components of the endomembrane system
- Recognize the relationship between the endomembrane system and its functions

The endomembrane system (endo = “within”) is a group of membranes and organelles ([\[link\]](#)) in eukaryotic cells that works together to modify, package, and transport lipids and proteins. It includes the nuclear envelope, lysosomes, and vesicles, which we’ve already mentioned, and the endoplasmic reticulum and Golgi apparatus, which we will cover shortly. Although not technically *within* the cell, the plasma membrane is included in the endomembrane system because, as you will see, it interacts with the other endomembranous organelles. The endomembrane system does not include the membranes of either mitochondria or chloroplasts.

### **Note:**

Art Connection



"Membrane and secretory proteins are synthesized in the rough endoplasmic reticulum (RER). The RER also sometimes modifies proteins. In this illustration, a (green) integral membrane protein in the ER is modified by attachment of a (purple) carbohydrate. Vesicles with the integral protein bud from the ER and fuse with the cis face of the Golgi apparatus. As the protein passes along the Golgi's cisternae, it is further modified by the addition of more carbohydrates. After its synthesis is

complete, it exits as integral membrane protein of the vesicle that bud from the Golgi's **trans** face and when the vesicle fuses with the cell membrane the protein becomes integral portion of that cell membrane.  
(credit: modification of work by Magnus Manske)

If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

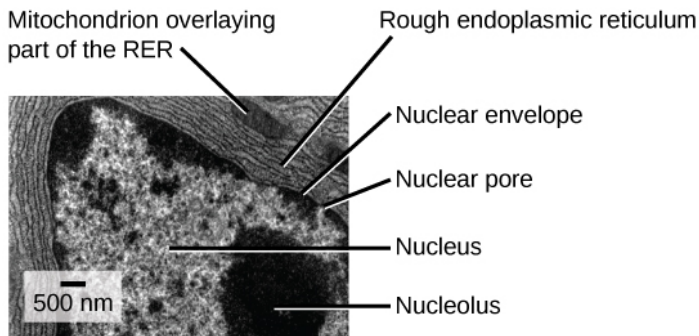
## The Endoplasmic Reticulum

The **endoplasmic reticulum (ER)** ([\[link\]](#)) is a series of interconnected membranous sacs and tubules that collectively modifies proteins and synthesizes lipids. However, these two functions are performed in separate areas of the ER: the rough ER and the smooth ER, respectively.

The hollow portion of the ER tubules is called the lumen or cisternal space. The membrane of the ER, which is a phospholipid bilayer embedded with proteins, is continuous with the nuclear envelope.

### Rough ER

The **rough endoplasmic reticulum (RER)** is so named because the ribosomes attached to its cytoplasmic surface give it a studded appearance when viewed through an electron microscope ([\[link\]](#)).



This transmission electron micrograph shows the rough endoplasmic reticulum and other organelles in a pancreatic cell. (credit: modification of work by Louisa Howard)

Ribosomes transfer their newly synthesized proteins into the lumen of the RER where they undergo structural modifications, such as folding or the acquisition of side chains. These modified proteins will be incorporated into cellular membranes—the membrane of the ER or those of other organelles—or secreted from the cell (such as protein hormones, enzymes). The RER also makes phospholipids for cellular membranes.

If the phospholipids or modified proteins are not destined to stay in the RER, they will reach their destinations via transport vesicles that bud from the RER's membrane ([\[link\]](#)).

Since the RER is engaged in modifying proteins (such as enzymes, for example) that will be secreted from the cell, you would be correct in assuming that the RER is abundant in cells that secrete proteins. This is the case with cells of the liver, for example.

## Smooth ER

The **smooth endoplasmic reticulum (SER)** is continuous with the RER but has few or no ribosomes on its cytoplasmic surface ([\[link\]](#)). Functions

of the SER include synthesis of carbohydrates, lipids, and steroid hormones; detoxification of medications and poisons; and storage of calcium ions.

In muscle cells, a specialized SER called the sarcoplasmic reticulum is responsible for storage of the calcium ions that are needed to trigger the coordinated contractions of the muscle cells.

**Note:**

Link to Learning



You can watch an excellent animation of the endomembrane system [here](#). At the end of the animation, there is a short self-assessment.

**Note:**

Career Connection

**Cardiologist**

Heart disease is the leading cause of death in the United States. This is primarily due to our sedentary lifestyle and our high trans-fat diets.

Heart failure is just one of many disabling heart conditions. Heart failure does not mean that the heart has stopped working. Rather, it means that the heart can't pump with sufficient force to transport oxygenated blood to all the vital organs. Left untreated, heart failure can lead to kidney failure and failure of other organs.

The wall of the heart is composed of cardiac muscle tissue. Heart failure occurs when the endoplasmic reticula of cardiac muscle cells do not

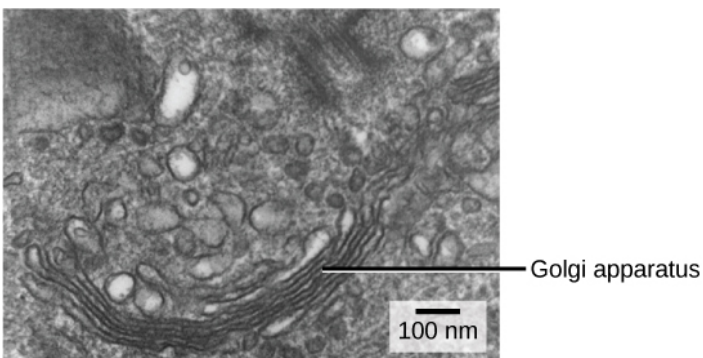
function properly. As a result, an insufficient number of calcium ions are available to trigger a sufficient contractile force.

Cardiologists (cardi- = “heart”; -ologist = “one who studies”) are doctors who specialize in treating heart diseases, including heart failure.

Cardiologists can make a diagnosis of heart failure via physical examination, results from an electrocardiogram (ECG, a test that measures the electrical activity of the heart), a chest X-ray to see whether the heart is enlarged, and other tests. If heart failure is diagnosed, the cardiologist will typically prescribe appropriate medications and recommend a reduction in table salt intake and a supervised exercise program.

## The Golgi Apparatus

We have already mentioned that vesicles can bud from the ER and transport their contents elsewhere, but where do the vesicles go? Before reaching their final destination, the lipids or proteins within the transport vesicles still need to be sorted, packaged, and tagged so that they wind up in the right place. Sorting, tagging, packaging, and distribution of lipids and proteins takes place in the **Golgi apparatus** (also called the Golgi body), a series of flattened membranes ([\[link\]](#)).



The Golgi apparatus in this white blood cell is visible as a stack of semicircular, flattened rings in the lower portion of the image. Several

vesicles can be seen near the Golgi apparatus. (credit: modification of work by Louisa Howard)

The receiving side of the Golgi apparatus is called the *cis* face. The opposite side is called the *trans* face. The transport vesicles that formed from the ER travel to the *cis* face, fuse with it, and empty their contents into the lumen of the Golgi apparatus. As the proteins and lipids travel through the Golgi, they undergo further modifications that allow them to be sorted. The most frequent modification is the addition of short chains of sugar molecules. These newly modified proteins and lipids are then tagged with phosphate groups or other small molecules so that they can be routed to their proper destinations.

Finally, the modified and tagged proteins are packaged into secretory vesicles that bud from the *trans* face of the Golgi. While some of these vesicles deposit their contents into other parts of the cell where they will be used, other secretory vesicles fuse with the plasma membrane and release their contents outside the cell.

In another example of form following function, cells that engage in a great deal of secretory activity (such as cells of the salivary glands that secrete digestive enzymes or cells of the immune system that secrete antibodies) have an abundance of Golgi.

In plant cells, the Golgi apparatus has the additional role of synthesizing polysaccharides, some of which are incorporated into the cell wall and some of which are used in other parts of the cell.

**Note:****Career Connection****Geneticist**

Many diseases arise from genetic mutations that prevent the synthesis of critical proteins. One such disease is Lowe disease (also called

oculocerebrorenal syndrome, because it affects the eyes, brain, and kidneys). In Lowe disease, there is a deficiency in an enzyme localized to the Golgi apparatus. Children with Lowe disease are born with cataracts, typically develop kidney disease after the first year of life, and may have impaired mental abilities.

Lowe disease is a genetic disease caused by a mutation on the X chromosome. The X chromosome is one of the two human sex chromosomes, as these chromosomes determine a person's sex. Females possess two X chromosomes while males possess one X and one Y chromosome. In females, the genes on only one of the two X chromosomes are expressed. Therefore, females who carry the Lowe disease gene on one of their X chromosomes have a 50/50 chance of having the disease.

However, males only have one X chromosome and the genes on this chromosome are always expressed. Therefore, males will always have Lowe disease if their X chromosome carries the Lowe disease gene. The location of the mutated gene, as well as the locations of many other mutations that cause genetic diseases, has now been identified. Through prenatal testing, a woman can find out if the fetus she is carrying may be afflicted with one of several genetic diseases.

Geneticists analyze the results of prenatal genetic tests and may counsel pregnant women on available options. They may also conduct genetic research that leads to new drugs or foods, or perform DNA analyses that are used in forensic investigations.

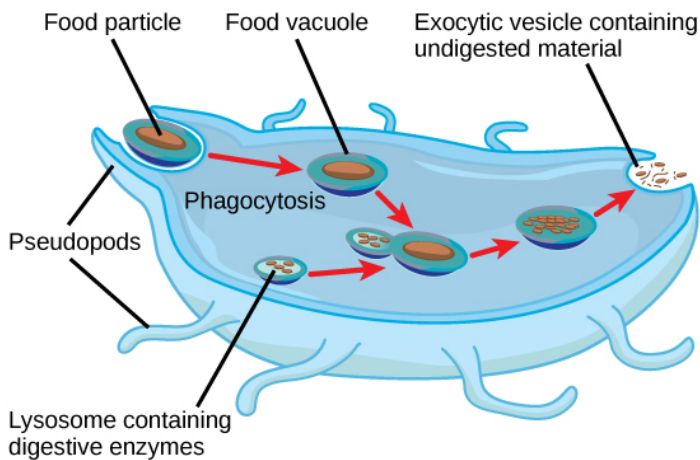
## **Lysosomes**

In addition to their role as the digestive component and organelle-recycling facility of animal cells, lysosomes are considered to be parts of the endomembrane system. Lysosomes also use their hydrolytic enzymes to destroy pathogens (disease-causing organisms) that might enter the cell. A good example of this occurs in a group of white blood cells called macrophages, which are part of your body's immune system. In a process known as phagocytosis or endocytosis, a section of the plasma membrane of the macrophage invaginates (folds in) and engulfs a pathogen. The invaginated section, with the pathogen inside, then pinches itself off from



the plasma membrane and becomes a vesicle. The vesicle fuses with a lysosome. The lysosome's hydrolytic enzymes then destroy the pathogen ([link](#)).

### Phagocytosis



A macrophage has engulfed (phagocytized) a potentially pathogenic bacterium and then fuses with a lysosomes within the cell to destroy the pathogen. Other organelles are present in the cell but for simplicity are not shown.

## Section Summary

The endomembrane system includes the nuclear envelope, lysosomes, vesicles, the ER, and Golgi apparatus, as well as the plasma membrane. These cellular components work together to modify, package, tag, and transport proteins and lipids that form the membranes.

The RER modifies proteins and synthesizes phospholipids used in cell membranes. The SER synthesizes carbohydrates, lipids, and steroid

hormones; engages in the detoxification of medications and poisons; and stores calcium ions. Sorting, tagging, packaging, and distribution of lipids and proteins take place in the Golgi apparatus. Lysosomes are created by the budding of the membranes of the RER and Golgi. Lysosomes digest macromolecules, recycle worn-out organelles, and destroy pathogens.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

---

#### Solution:

[\[link\]](#) It would end up on the outside. After the vesicle passes through the Golgi apparatus and fuses with the plasma membrane, it turns inside out.

## Review Questions

### Exercise:

#### Problem:

Which of the following is not a component of the endomembrane system?

- a. mitochondrion
  - b. Golgi apparatus
  - c. endoplasmic reticulum
  - d. lysosome
- 

#### Solution:

A

**Exercise:**

**Problem:**

The process by which a cell engulfs a foreign particle is known as:

- a. endosymbiosis
- b. phagocytosis
- c. hydrolysis
- d. membrane synthesis

---

**Solution:**

B

**Exercise:**

**Problem:**

Which of the following is most likely to have the greatest concentration of smooth endoplasmic reticulum?

- a. a cell that secretes enzymes
- b. a cell that destroys pathogens
- c. a cell that makes steroid hormones
- d. a cell that engages in photosynthesis

---

**Solution:**

C

**Exercise:**

**Problem:**

Which of the following sequences correctly lists in order the steps involved in the incorporation of a proteinaceous molecule within a cell?

- a. synthesis of the protein on the ribosome; modification in the Golgi apparatus; packaging in the endoplasmic reticulum; tagging in the vesicle
  - b. synthesis of the protein on the lysosome; tagging in the Golgi; packaging in the vesicle; distribution in the endoplasmic reticulum
  - c. synthesis of the protein on the ribosome; modification in the endoplasmic reticulum; tagging in the Golgi; distribution via the vesicle
  - d. synthesis of the protein on the lysosome; packaging in the vesicle; distribution via the Golgi; tagging in the endoplasmic reticulum
- 

**Solution:**

C

## Free Response

**Exercise:**

**Problem:**

In the context of cell biology, what do we mean by form follows function? What are at least two examples of this concept?

---

**Solution:**

“Form follows function” refers to the idea that the function of a body part dictates the form of that body part. As an example, compare your arm to a bat’s wing. While the bones of the two correspond, the parts serve different functions in each organism and their forms have adapted to follow that function.

**Exercise:**

**Problem:**

In your opinion, is the nuclear membrane part of the endomembrane system? Why or why not? Defend your answer.

---

**Solution:**

Since the external surface of the nuclear membrane is continuous with the rough endoplasmic reticulum, which is part of the endomembrane system, then it is correct to say that it is part of the system.

**Glossary**

endomembrane system

group of organelles and membranes in eukaryotic cells that work together modifying, packaging, and transporting lipids and proteins

endoplasmic reticulum (ER)

series of interconnected membranous structures within eukaryotic cells that collectively modify proteins and synthesize lipids

Golgi apparatus

eukaryotic organelle made up of a series of stacked membranes that sorts, tags, and packages lipids and proteins for distribution

rough endoplasmic reticulum (RER)

region of the endoplasmic reticulum that is studded with ribosomes and engages in protein modification and phospholipid synthesis

smooth endoplasmic reticulum (SER)

region of the endoplasmic reticulum that has few or no ribosomes on its cytoplasmic surface and synthesizes carbohydrates, lipids, and steroid hormones; detoxifies certain chemicals (like pesticides, preservatives, medications, and environmental pollutants), and stores calcium ions

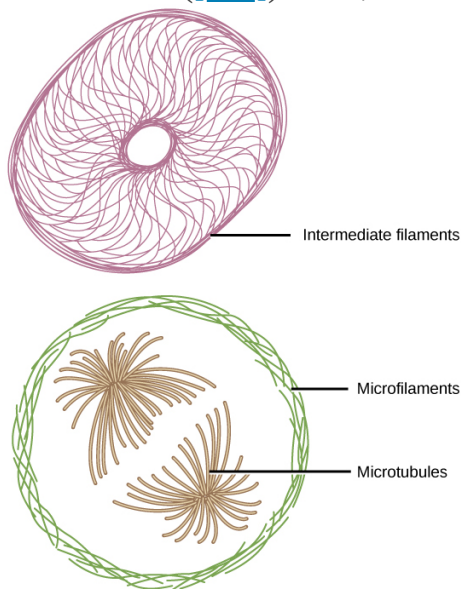
## The Cytoskeleton

By the end of this section, you will be able to:

- Describe the cytoskeleton
- Compare the roles of microfilaments, intermediate filaments, and microtubules
- Compare and contrast cilia and flagella
- Summarize the differences among the components of prokaryotic cells, animal cells, and plant cells

If you were to remove all the organelles from a cell, would the plasma membrane and the cytoplasm be the only components left? No. Within the cytoplasm, there would still be ions and organic molecules, plus a network of protein fibers that help maintain the shape of the cell, secure some organelles in specific positions, allow cytoplasm and vesicles to move within the cell, and enable cells within multicellular organisms to move.

Collectively, this network of protein fibers is known as the **cytoskeleton**. There are three types of fibers within the cytoskeleton: microfilaments, intermediate filaments, and microtubules ([\[link\]](#)). Here, we will examine each.



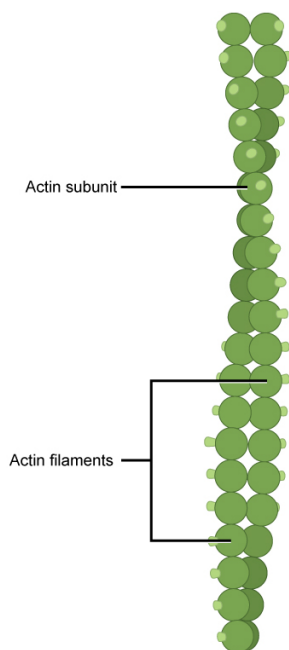
Microfilaments thicken the cortex around the inner edge of a cell; like rubber bands, they resist tension.

Microtubules are found in the interior of the cell where they maintain cell shape by resisting compressive forces.

Intermediate filaments are found throughout the cell and hold organelles in place.

## Microfilaments

Of the three types of protein fibers in the cytoskeleton, **microfilaments** are the narrowest. They function in cellular movement, have a diameter of about 7 nm, and are made of two intertwined strands of a globular protein called actin ([link](#)). For this reason, microfilaments are also known as actin filaments.



Microfilaments are made of two intertwined strands of actin.

Actin is powered by ATP to assemble its filamentous form, which serves as a track for the movement of a motor protein called myosin. This enables actin to engage in cellular events requiring motion, such as cell division in animal cells and cytoplasmic streaming, which is the circular movement of the cell cytoplasm in plant cells. Actin and myosin are plentiful in muscle cells. When your actin and myosin filaments slide past each other, your muscles contract.

Microfilaments also provide some rigidity and shape to the cell. They can depolymerize (disassemble) and reform quickly, thus enabling a cell to change its shape and move.

White blood cells (your body's infection-fighting cells) make good use of this ability. They can move to the site of an infection and phagocytize the pathogen.

**Note:**

Link to Learning



To see an example of a white blood cell in action, watch a short time-lapse video of the cell capturing two bacteria. It engulfs one and then moves on to the other.

[https://www.openstaxcollege.org/l/chasing\\_bacteria](https://www.openstaxcollege.org/l/chasing_bacteria)

## Intermediate Filaments

Intermediate filaments are made of several strands of fibrous proteins that are wound together ([link]). These elements of the cytoskeleton get their name from the fact that their diameter, 8 to 10 nm, is between those of microfilaments and microtubules.



Intermediate filaments consist of several intertwined strands of fibrous proteins.

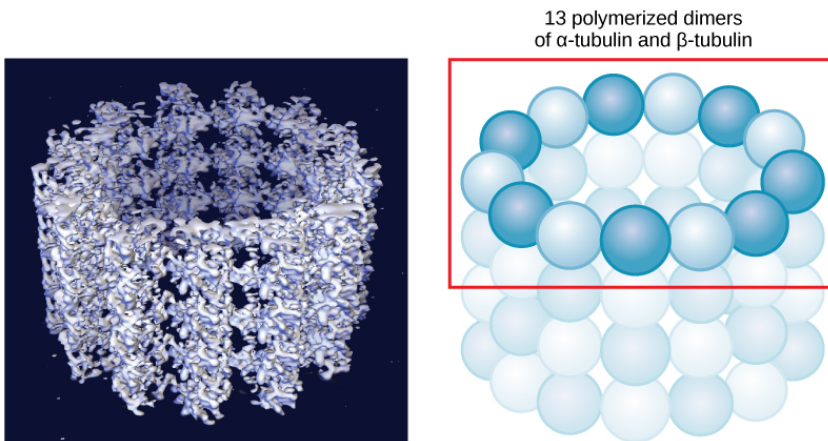
**Intermediate filaments** have no role in cell movement. Their function is purely structural. They bear tension, thus maintaining the shape of the cell, and anchor the nucleus and other organelles in place. [link] shows how intermediate filaments create a supportive scaffolding inside the cell.



The intermediate filaments are the most diverse group of cytoskeletal elements. Several types of fibrous proteins are found in the intermediate filaments. You are probably most familiar with keratin, the fibrous protein that strengthens your hair, nails, and the epidermis of the skin.

## Microtubules

As their name implies, microtubules are small hollow tubes. The walls of the microtubule are made of polymerized dimers of  $\alpha$ -tubulin and  $\beta$ -tubulin, two globular proteins ([\[link\]](#)). With a diameter of about 25 nm, **microtubules** are the widest components of the cytoskeleton. They help the cell resist compression, provide a track along which vesicles move through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. Like microfilaments, microtubules can dissolve and reform quickly.



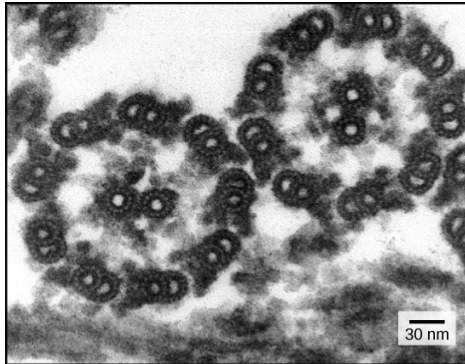
Microtubules are hollow. Their walls consist of 13 polymerized dimers of  $\alpha$ -tubulin and  $\beta$ -tubulin (right image). The left image shows the molecular structure of the tube.

Microtubules are also the structural elements of flagella, cilia, and centrioles (the latter are the two perpendicular bodies of the centrosome). In fact, in animal cells, the centrosome is the microtubule-organizing center. In eukaryotic cells, flagella and cilia are quite different structurally from their counterparts in prokaryotes, as discussed below.

## Flagella and Cilia

To refresh your memory, **flagella** (singular = flagellum) are long, hair-like structures that extend from the plasma membrane and are used to move an entire cell (for example, sperm, *Euglena*). When present, the cell has just one flagellum or a few flagella. When **cilia** (singular = cilium) are present, however, many of them extend along the entire surface of the plasma membrane. They are short, hair-like structures that are used to move entire cells (such as paramecia) or substances along the outer surface of the cell (for example, the cilia of cells lining the Fallopian tubes that move the ovum toward the uterus, or cilia lining the cells of the respiratory tract that trap particulate matter and move it toward your nostrils.)

Despite their differences in length and number, flagella and cilia share a common structural arrangement of microtubules called a “9 + 2 array.” This is an appropriate name because a single flagellum or cilium is made of a ring of nine microtubule doublets, surrounding a single microtubule doublet in the center ([\[link\]](#)).



This transmission electron micrograph of two flagella shows the 9 + 2 array of microtubules: nine microtubule doublets surround a single microtubule doublet. (credit: modification of work by Dartmouth Electron Microscope Facility, Dartmouth College; scale-bar data from Matt Russell)

You have now completed a broad survey of the components of prokaryotic and eukaryotic cells. For a summary of cellular components in prokaryotic and eukaryotic cells, see [\[link\]](#).

<b>Components of Prokaryotic and Eukaryotic Cells</b>				
<b>Cell Component</b>	<b>Function</b>	<b>Present in Prokaryotes?</b>	<b>Present in Animal Cells?</b>	<b>Present in Plant Cells?</b>
Plasma membrane	Separates cell from external environment; controls passage of organic molecules, ions, water, oxygen, and wastes into and out of cell	Yes	Yes	Yes
Cytoplasm	Provides turgor pressure to plant cells as fluid inside the central vacuole; site of many metabolic reactions; medium in which organelles are found	Yes	Yes	Yes
Nucleolus	Darkened area within the nucleus where ribosomal subunits are synthesized.	No	Yes	Yes
Nucleus	Cell organelle that houses DNA and directs synthesis of ribosomes and proteins	No	Yes	Yes
Ribosomes	Protein synthesis	Yes	Yes	Yes

<b>Components of Prokaryotic and Eukaryotic Cells</b>				
<b>Cell Component</b>	<b>Function</b>	<b>Present in Prokaryotes?</b>	<b>Present in Animal Cells?</b>	<b>Present in Plant Cells?</b>
Mitochondria	ATP production/cellular respiration	No	Yes	Yes
Peroxisomes	Oxidizes and thus breaks down fatty acids and amino acids, and detoxifies poisons	No	Yes	Yes
Vesicles and vacuoles	Storage and transport; digestive function in plant cells	No	Yes	Yes
Centrosome	Unspecified role in cell division in animal cells; source of microtubules in animal cells	No	Yes	No
Lysosomes	Digestion of macromolecules; recycling of worn-out organelles	No	Yes	No
Cell wall	Protection, structural support and maintenance of cell shape	Yes, primarily peptidoglycan	No	Yes, primarily cellulose
Chloroplasts	Photosynthesis	No	No	Yes

<b>Components of Prokaryotic and Eukaryotic Cells</b>				
<b>Cell Component</b>	<b>Function</b>	<b>Present in Prokaryotes?</b>	<b>Present in Animal Cells?</b>	<b>Present in Plant Cells?</b>
Endoplasmic reticulum	Modifies proteins and synthesizes lipids	No	Yes	Yes
Golgi apparatus	Modifies, sorts, tags, packages, and distributes lipids and proteins	No	Yes	Yes
Cytoskeleton	Maintains cell's shape, secures organelles in specific positions, allows cytoplasm and vesicles to move within cell, and enables unicellular organisms to move independently	Yes	Yes	Yes
Flagella	Cellular locomotion	Some	Some	No, except for some plant sperm cells.
Cilia	Cellular locomotion, movement of particles along extracellular surface of plasma membrane, and filtration	Some	Some	No

## Section Summary

The cytoskeleton has three different types of protein elements. From narrowest to widest, they are the microfilaments (actin filaments), intermediate filaments, and microtubules. Microfilaments are often associated with myosin. They provide rigidity and shape to the cell and facilitate cellular movements. Intermediate filaments bear tension and anchor the nucleus and other organelles in place. Microtubules help the cell resist compression, serve as tracks for motor proteins that move vesicles through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. They are also the structural element of centrioles, flagella, and cilia.

## Review Questions

### Exercise:

#### Problem:

Which of the following have the ability to disassemble and reform quickly?

- a. microfilaments and intermediate filaments
- b. microfilaments and microtubules
- c. intermediate filaments and microtubules
- d. only intermediate filaments

---

#### Solution:

B

### Exercise:

**Problem:** Which of the following do not play a role in intracellular movement?

- a. microfilaments and intermediate filaments
- b. microfilaments and microtubules
- c. intermediate filaments and microtubules
- d. only intermediate filaments

---

#### Solution:

D

## Free Response

**Exercise:****Problem:**

What are the similarities and differences between the structures of centrioles and flagella?

---

**Solution:**

Centrioles and flagella are alike in that they are made up of microtubules. In centrioles, two rings of nine microtubule “triplets” are arranged at right angles to one another. This arrangement does not occur in flagella.

**Exercise:**

**Problem:**How do cilia and flagella differ?

---

**Solution:**

Cilia and flagella are alike in that they are made up of microtubules. Cilia are short, hair-like structures that exist in large numbers and usually cover the entire surface of the plasma membrane. Flagella, in contrast, are long, hair-like structures; when flagella are present, a cell has just one or two.

**Glossary****cilium**

(plural = cilia) short, hair-like structure that extends from the plasma membrane in large numbers and is used to move an entire cell or move substances along the outer surface of the cell

**cytoskeleton**

network of protein fibers that collectively maintain the shape of the cell, secure some organelles in specific positions, allow cytoplasm and vesicles to move within the cell, and enable unicellular organisms to move independently

**flagellum**

(plural = flagella) long, hair-like structure that extends from the plasma membrane and is used to move the cell

**intermediate filament**

cytoskeletal component, composed of several intertwined strands of fibrous protein, that bears tension, supports cell-cell junctions, and anchors cells to extracellular structures

microfilament

narrowest element of the cytoskeleton system; it provides rigidity and shape to the cell and enables cellular movements

microtubule

widest element of the cytoskeleton system; it helps the cell resist compression, provides a track along which vesicles move through the cell, pulls replicated chromosomes to opposite ends of a dividing cell, and is the structural element of centrioles, flagella, and cilia



## Connections between Cells and Cellular Activities

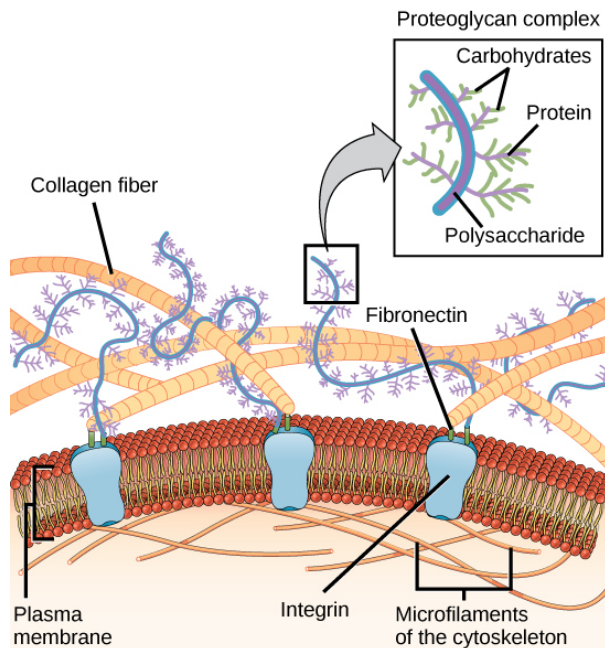
By the end of this section, you will be able to:

- Describe the extracellular matrix
- List examples of the ways that plant cells and animal cells communicate with adjacent cells
- Summarize the roles of tight junctions, desmosomes, gap junctions, and plasmodesmata

You already know that a group of similar cells working together is called a tissue. As you might expect, if cells are to work together, they must communicate with each other, just as you need to communicate with others if you work on a group project. Let's take a look at how cells communicate with each other.

### Extracellular Matrix of Animal Cells

Most animal cells release materials into the extracellular space. The primary components of these materials are proteins, and the most abundant protein is collagen. Collagen fibers are interwoven with carbohydrate-containing protein molecules called proteoglycans. Collectively, these materials are called the **extracellular matrix** ([\[link\]](#)). Not only does the extracellular matrix hold the cells together to form a tissue, but it also allows the cells within the tissue to communicate with each other. How can this happen?



The extracellular matrix consists of a network of proteins and carbohydrates.

Cells have protein receptors on the extracellular surfaces of their plasma membranes. When a molecule within the matrix binds to the receptor, it changes the molecular structure of the receptor. The receptor, in turn, changes the conformation of the microfilaments positioned just inside the plasma membrane. These conformational changes induce chemical signals inside the cell that reach the nucleus and turn “on” or “off” the transcription of specific sections of DNA, which affects the production of associated proteins, thus changing the activities within the cell.

Blood clotting provides an example of the role of the extracellular matrix in cell communication. When the cells lining a blood vessel are damaged, they display a protein receptor called tissue factor. When tissue factor binds with another factor in the extracellular matrix, it causes platelets to adhere to the wall of the damaged blood vessel, stimulates the adjacent smooth muscle cells in the blood vessel to contract (thus constricting the blood vessel), and

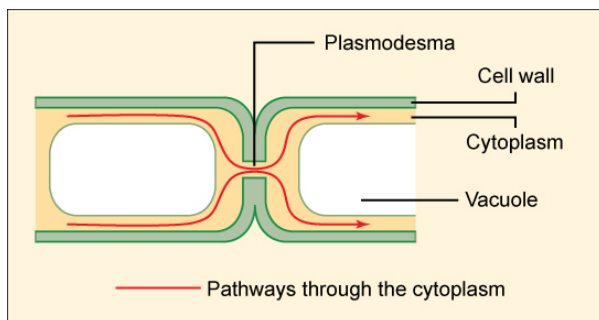
initiates a series of steps that stimulate the platelets to produce clotting factors.

## Intercellular Junctions

Cells can also communicate with each other via direct contact, referred to as intercellular junctions. There are some differences in the ways that plant and animal cells do this. Plasmodesmata are junctions between plant cells, whereas animal cell contacts include tight junctions, gap junctions, and desmosomes.

### Plasmodesmata

In general, long stretches of the plasma membranes of neighboring plant cells cannot touch one another because they are separated by the cell wall that surrounds each cell ([link](#))b). How then, can a plant transfer water and other soil nutrients from its roots, through its stems, and to its leaves? Such transport uses the vascular tissues (xylem and phloem) primarily. There also exist structural modifications called **plasmodesmata** (singular = plasmodesma), numerous channels that pass between cell walls of adjacent plant cells, connect their cytoplasm, and enable materials to be transported from cell to cell, and thus throughout the plant ([link](#)).

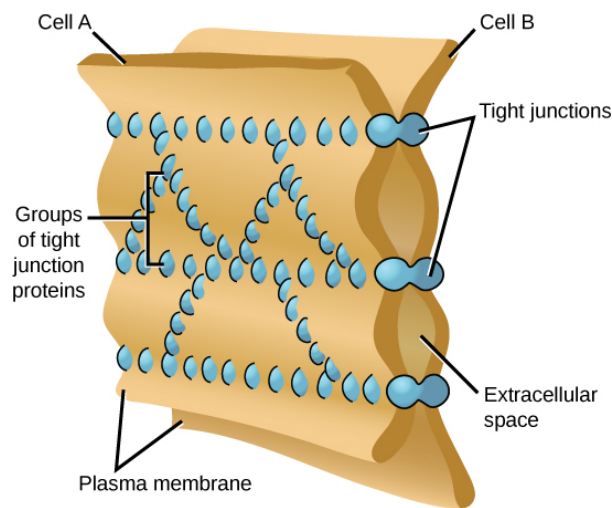


A plasmodesma is a channel between the cell walls of two adjacent plant cells.

Plasmodesmata allow materials to pass from the cytoplasm of one plant cell to the cytoplasm of an adjacent cell.

## Tight Junctions

A **tight junction** is a watertight seal between two adjacent animal cells ([link](#)). The cells are held tightly against each other by proteins (predominantly two proteins called claudins and occludins).



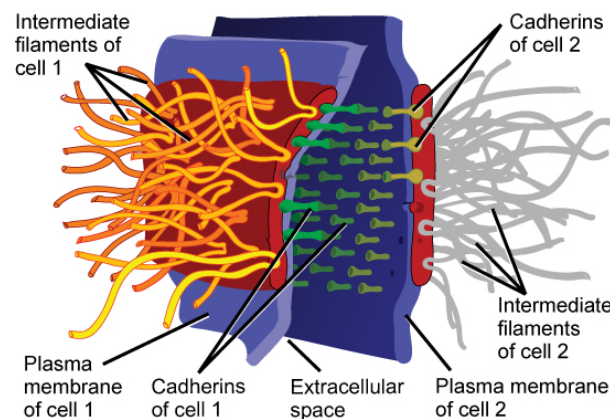
Tight junctions form watertight connections between adjacent animal cells. Proteins create tight junction adherence.

(credit: modification of work by Mariana Ruiz Villareal)

This tight adherence prevents materials from leaking between the cells; tight junctions are typically found in epithelial tissues that line internal organs and cavities, and comprise most of the skin. For example, the tight junctions of the epithelial cells lining your urinary bladder prevent urine from leaking out into the extracellular space.

## Desmosomes

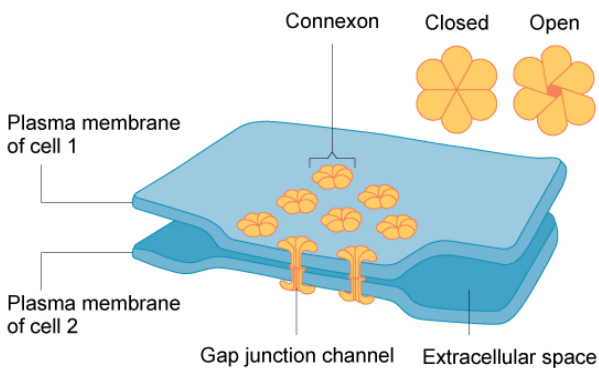
Also found only in animal cells are **desmosomes**, which act like spot welds between adjacent epithelial cells ([\[link\]](#)). Short proteins called cadherins in the plasma membrane connect to intermediate filaments to create desmosomes. The cadherins join two adjacent cells together and maintain the cells in a sheet-like formation in organs and tissues that stretch, like the skin, heart, and muscles.



A desmosome forms a very strong spot weld between cells. It is created by the linkage of cadherins and intermediate filaments. (credit: modification of work by Mariana Ruiz Villareal)

## Gap Junctions

**Gap junctions** in animal cells are like plasmodesmata in plant cells in that they are channels between adjacent cells that allow for the transport of ions, nutrients, and other substances that enable cells to communicate ([\[link\]](#)). Structurally, however, gap junctions and plasmodesmata differ.



A gap junction is a protein-lined pore that allows water and small molecules to pass between adjacent animal cells. (credit: modification of work by Mariana Ruiz Villareal)

Gap junctions develop when a set of six proteins (called connexins) in the plasma membrane arrange themselves in an elongated donut-like configuration called a connexon. When the pores (“doughnut holes”) of connexons in adjacent animal cells align, a channel between the two cells forms. Gap junctions are particularly important in cardiac muscle: The electrical signal for the muscle to contract is passed efficiently through gap junctions, allowing the heart muscle cells to contract in tandem.

**Note:**

## Link to Learning



To conduct a virtual microscopy lab and review the parts of a cell, work through the steps of this [interactive assignment](#).

## Section Summary

Animal cells communicate via their extracellular matrices and are connected to each other via tight junctions, desmosomes, and gap junctions. Plant cells are connected and communicate with each other via plasmodesmata.

When protein receptors on the surface of the plasma membrane of an animal cell bind to a substance in the extracellular matrix, a chain of reactions begins that changes activities taking place within the cell. Plasmodesmata are channels between adjacent plant cells, while gap junctions are channels between adjacent animal cells. However, their structures are quite different. A tight junction is a watertight seal between two adjacent cells, while a desmosome acts like a spot weld.

## Review Questions

### Exercise:

**Problem:** Which of the following are found only in plant cells?

- a. gap junctions
- b. desmosomes
- c. plasmodesmata

d. tight junctions

---

**Solution:**

C

**Exercise:**

**Problem:**

The key components of desmosomes are cadherins and \_\_\_\_\_.

- a. actin
- b. microfilaments
- c. intermediate filaments
- d. microtubules

---

**Solution:**

C

**Free Response**

**Exercise:**

**Problem:**

How does the structure of a plasmodesma differ from that of a gap junction?

---

**Solution:**

They differ because plant cell walls are rigid. Plasmodesmata, which a plant cell needs for transportation and communication, are able to allow movement of really large molecules. Gap junctions are necessary in animal cells for transportation and communication.

**Exercise:**



**Problem:** Explain how the extracellular matrix functions.

---

**Solution:**

The extracellular matrix functions in support and attachment for animal tissues. It also functions in the healing and growth of the tissue.

## Glossary

desmosome

linkages between adjacent epithelial cells that form when cadherins in the plasma membrane attach to intermediate filaments

extracellular matrix

material (primarily collagen, glycoproteins, and proteoglycans) secreted from animal cells that provides mechanical protection and anchoring for the cells in the tissue

gap junction

channel between two adjacent animal cells that allows ions, nutrients, and low molecular weight substances to pass between cells, enabling the cells to communicate

plasmodesma

(plural = plasmodesmata) channel that passes between the cell walls of adjacent plant cells, connects their cytoplasm, and allows materials to be transported from cell to cell

tight junction

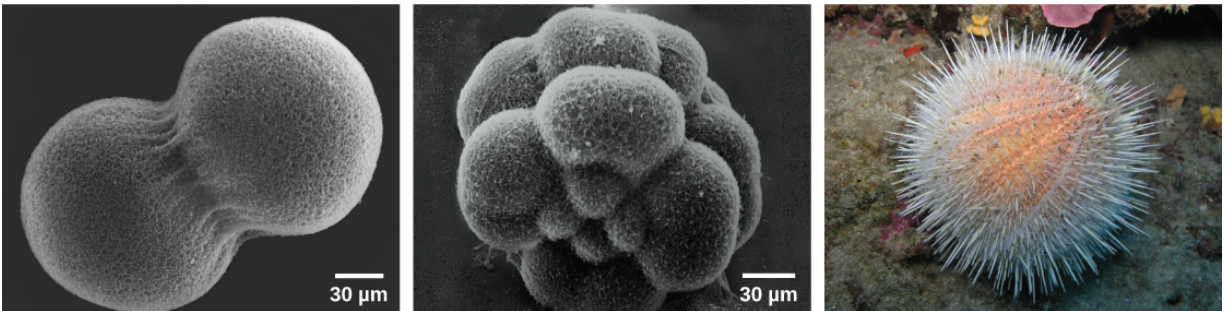
firm seal between two adjacent animal cells created by protein adherence

## Introduction

class="introduction"

A sea urchin begins life as a single cell that (a) divides to form two cells, visible by scanning electron microscopy. After four rounds of cell division, (b) there are 16 cells, as seen in this SEM image. After many rounds of cell division, the individual develops into a complex, multicellular organism, as seen in this (c) mature sea urchin. (credit a: modificatio

n of work  
by Evelyn  
Spiegel,  
Louisa  
Howard;  
credit b:  
modificatio  
n of work  
by Evelyn  
Spiegel,  
Louisa  
Howard;  
credit c:  
modificatio  
n of work  
by Marco  
Busdraghi;  
scale-bar  
data from  
Matt  
Russell)



A human, as well as every sexually reproducing organism, begins life as a fertilized egg (embryo) or zygote. Trillions of cell divisions subsequently occur in a controlled manner to produce a complex, multicellular human. In other words, that original single cell is the ancestor of every other cell in the body. Once a being is fully grown, cell reproduction is still necessary to repair or regenerate tissues. For example, new blood and skin cells are

constantly being produced. All multicellular organisms use cell division for growth and the maintenance and repair of cells and tissues. Cell division is tightly regulated, and the occasional failure of regulation can have life-threatening consequences. Single-celled organisms use cell division as their method of reproduction.

## Cell Division

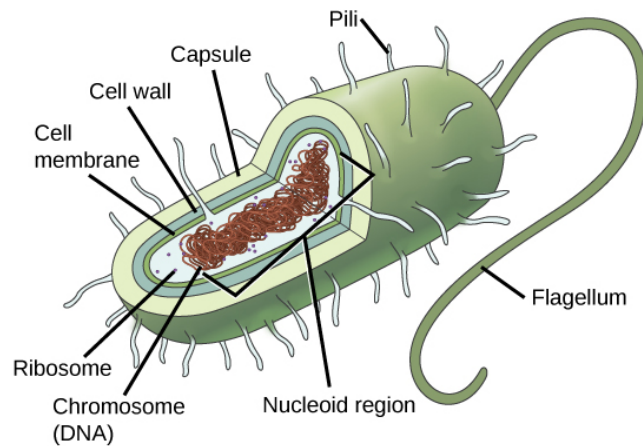
By the end of this section, you will be able to:

- Describe the structure of prokaryotic and eukaryotic genomes
- Distinguish between chromosomes, genes, and traits
- Describe the mechanisms of chromosome compaction

The continuity of life from one cell to another has its foundation in the reproduction of cells by way of the cell cycle. The **cell cycle** is an orderly sequence of events that describes the stages of a cell's life from the division of a single parent cell to the production of two new daughter cells. The mechanisms involved in the cell cycle are highly regulated.

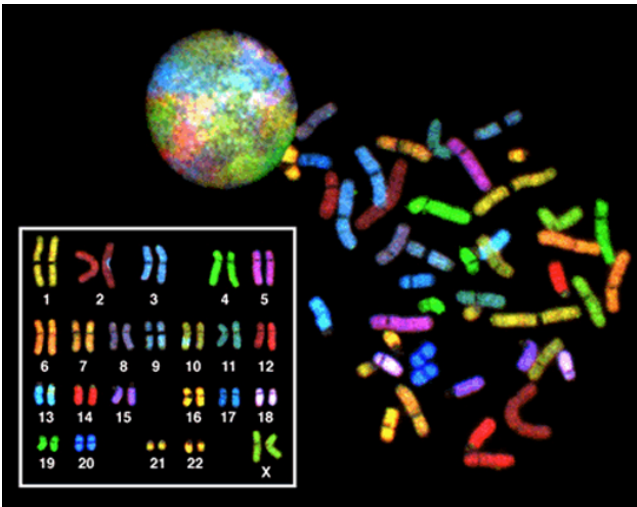
## Genomic DNA

Before discussing the steps a cell must undertake to replicate, a deeper understanding of the structure and function of a cell's genetic information is necessary. A cell's DNA, packaged as a double-stranded DNA molecule, is called its **genome**. In prokaryotes, the genome is composed of a single, double-stranded DNA molecule in the form of a loop or circle ([\[link\]](#)). The region in the cell containing this genetic material is called a nucleoid. Some prokaryotes also have smaller loops of DNA called plasmids that are not essential for normal growth. Bacteria can exchange these plasmids with other bacteria, sometimes receiving beneficial new genes that the recipient can add to their chromosomal DNA. Antibiotic resistance is one trait that often spreads through a bacterial colony through plasmid exchange.



Prokaryotes, including bacteria and archaea, have a single, circular chromosome located in a central region called the nucleoid.

In eukaryotes, the genome consists of several double-stranded linear DNA molecules ([\[link\]](#)). Each species of eukaryotes has a characteristic number of chromosomes in the nuclei of its cells. Human body cells have 46 chromosomes, while human **gametes** (sperm or eggs) have 23 chromosomes each. A typical body cell, or somatic cell, contains two matched sets of chromosomes, a configuration known as **diploid**. The letter  $n$  is used to represent a single set of chromosomes; therefore, a diploid organism is designated  $2n$ . Human cells that contain one set of chromosomes are called gametes, or sex cells; these are eggs and sperm, and are designated  $1n$ , or **haploid**.



There are 23 pairs of homologous chromosomes in a female human somatic cell. The condensed chromosomes are viewed within the nucleus (top), removed from a cell in mitosis and spread out on a slide (right), and artificially arranged according to length (left); an arrangement like this is called a karyotype. In this image, the chromosomes were exposed to fluorescent stains for differentiation of the different chromosomes. A method of staining called “chromosome painting” employs fluorescent dyes that highlight chromosomes in different colors. (credit: National Human Genome Project/NIH)

Matched pairs of chromosomes in a diploid organism are called **homologous** (“same knowledge”) **chromosomes**. Homologous

chromosomes are the same length and have specific nucleotide segments called **genes** in exactly the same location, or **locus**. Genes, the functional units of chromosomes, determine specific characteristics by coding for specific proteins. Traits are the variations of those characteristics. For example, hair color is a characteristic with traits that are blonde, brown, or black.

Each copy of a homologous pair of chromosomes originates from a different parent; therefore, the genes themselves are not identical. The variation of individuals within a species is due to the specific combination of the genes inherited from both parents. Even a slightly altered sequence of nucleotides within a gene can result in an alternative trait. For example, there are three possible gene sequences on the human chromosome that code for blood type: sequence A, sequence B, and sequence O. Because all diploid human cells have two copies of the chromosome that determines blood type, the blood type (the trait) is determined by which two versions of the marker gene are inherited. It is possible to have two copies of the same gene sequence on both homologous chromosomes, with one on each (for example, AA, BB, or OO), or two different sequences, such as AB.

Minor variations of traits, such as blood type, eye color, and handedness, contribute to the natural variation found within a species. However, if the entire DNA sequence from any pair of human homologous chromosomes is compared, the difference is less than one percent. The sex chromosomes, X and Y, are the single exception to the rule of homologous chromosome uniformity: Other than a small amount of homology that is necessary to accurately produce gametes, the genes found on the X and Y chromosomes are different.

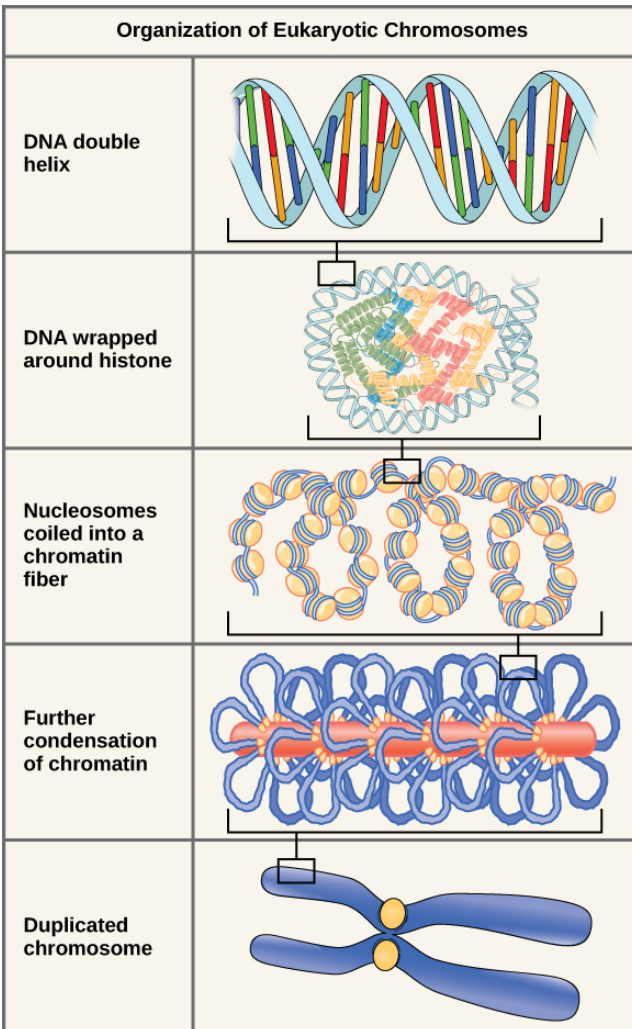
## **Eukaryotic Chromosomal Structure and Compaction**

If the DNA from all 46 chromosomes in a human cell nucleus was laid out end to end, it would measure approximately two meters; however, its diameter would be only 2 nm. Considering that the size of a typical human cell is about 10  $\mu\text{m}$  (100,000 cells lined up to equal one meter), DNA must be tightly packaged to fit in the cell's nucleus. At the same time, it must also be readily accessible for the genes to be expressed. During some stages



of the cell cycle, the long strands of DNA are condensed into compact chromosomes. There are a number of ways that chromosomes are compacted.

In the first level of compaction, short stretches of the DNA double helix wrap around a core of eight **histone proteins** at regular intervals along the entire length of the chromosome ([\[link\]](#)). The DNA-histone complex is called chromatin. The beadlike, histone DNA complex is called a **nucleosome**, and DNA connecting the nucleosomes is called linker DNA. A DNA molecule in this form is about seven times shorter than the double helix without the histones, and the beads are about 10 nm in diameter, in contrast with the 2-nm diameter of a DNA double helix. The next level of compaction occurs as the nucleosomes and the linker DNA between them are coiled into a 30-nm chromatin fiber. This coiling further shortens the chromosome so that it is now about 50 times shorter than the extended form. In the third level of packing, a variety of fibrous proteins is used to pack the chromatin. These fibrous proteins also ensure that each chromosome in a non-dividing cell occupies a particular area of the nucleus that does not overlap with that of any other chromosome (see the top image in [\[link\]](#)).



Double-stranded DNA wraps around histone proteins to form nucleosomes that have the appearance of “beads on a string.” The nucleosomes are coiled into a 30-nm chromatin fiber. When a cell undergoes mitosis, the chromosomes condense even further.

DNA replicates in the S phase of interphase. After replication, the chromosomes are composed of two linked sister **chromatids**. When fully

compact, the pairs of identically packed chromosomes are bound to each other by cohesin proteins. The connection between the sister chromatids is closest in a region called the **centromere**. The conjoined sister chromatids, with a diameter of about 1  $\mu\text{m}$ , are visible under a light microscope. The centromeric region is highly condensed and thus will appear as a constricted area.

**Note:**

Link to Learning



This animation illustrates the different levels of chromosome packing.  
[https://www.openstaxcollege.org/l/Packaged\\_DNA](https://www.openstaxcollege.org/l/Packaged_DNA)

## Section Summary

Prokaryotes have a single circular chromosome composed of double-stranded DNA, whereas eukaryotes have multiple, linear chromosomes composed of chromatin surrounded by a nuclear membrane. The 46 chromosomes of human somatic cells are composed of 22 pairs of autosomes (matched pairs) and a pair of sex chromosomes, which may or may not be matched. This is the  $2n$  or diploid state. Human gametes have 23 chromosomes or one complete set of chromosomes; a set of chromosomes is complete with either one of the sex chromosomes. This is the  $n$  or haploid state. Genes are segments of DNA that code for a specific protein. An organism's traits are determined by the genes inherited from each parent. Duplicated chromosomes are composed of two sister

chromatids. Chromosomes are compacted using a variety of mechanisms during certain stages of the cell cycle. Several classes of protein are involved in the organization and packing of the chromosomal DNA into a highly condensed structure. The condensing complex compacts chromosomes, and the resulting condensed structure is necessary for chromosomal segregation during mitosis.

## Review Questions

### Exercise:

#### Problem:

A diploid cell has \_\_\_\_\_ the number of chromosomes as a haploid cell.

- a. one-fourth
- b. half
- c. twice
- d. four times

---

#### Solution:

C

### Exercise:

#### Problem:

An organism's traits are determined by the specific combination of inherited \_\_\_\_\_.

- a. cells.
- b. genes.
- c. proteins.
- d. chromatids.

---

**Solution:**

B

**Exercise:**

**Problem:**

The first level of DNA organization in a eukaryotic cell is maintained by which molecule?

- a. cohesin
- b. condensin
- c. chromatin
- d. histone

---

**Solution:**

D

**Exercise:**

**Problem:**

Identical copies of chromatin held together by cohesin at the centromere are called \_\_\_\_\_.

- a. histones.
- b. nucleosomes.
- c. chromatin.
- d. sister chromatids.

---

**Solution:**

D

**Free Response**

**Exercise:****Problem:**

Compare and contrast a human somatic cell to a human gamete.

---

**Solution:**

Human somatic cells have 46 chromosomes: 22 pairs and 2 sex chromosomes that may or may not form a pair. This is the  $2n$  or diploid condition. Human gametes have 23 chromosomes, one each of 23 unique chromosomes, one of which is a sex chromosome. This is the  $n$  or haploid condition.

**Exercise:****Problem:**

What is the relationship between a genome, chromosomes, and genes?

---

**Solution:**

The genome consists of the sum total of an organism's chromosomes. Each chromosome contains hundreds and sometimes thousands of genes, segments of DNA that code for a polypeptide or RNA, and a large amount of DNA with no known function.

**Exercise:****Problem:**

Eukaryotic chromosomes are thousands of times longer than a typical cell. Explain how chromosomes can fit inside a eukaryotic nucleus.

---

**Solution:**

The DNA double helix is wrapped around histone proteins to form structures called nucleosomes. Nucleosomes and the linker DNA in between them are coiled into a 30-nm fiber. During cell division, chromatin is further condensed by packing proteins.

## Glossary

### cell cycle

ordered sequence of events that a cell passes through between one cell division and the next

### centromere

region at which sister chromatids are bound together; a constricted area in condensed chromosomes

### chromatid

single DNA molecule of two strands of duplicated DNA and associated proteins held together at the centromere

### diploid

cell, nucleus, or organism containing two sets of chromosomes ( $2n$ )

### gamete

haploid reproductive cell or sex cell (sperm, pollen grain, or egg)

### gene

physical and functional unit of heredity, a sequence of DNA that codes for a protein.

### genome

total genetic information of a cell or organism

### haploid

cell, nucleus, or organism containing one set of chromosomes ( $n$ )

### histone

one of several similar, highly conserved, low molecular weight, basic proteins found in the chromatin of all eukaryotic cells; associates with DNA to form nucleosomes

### homologous chromosomes

chromosomes of the same morphology with genes in the same location; diploid organisms have pairs of homologous chromosomes

(homologs), with each homolog derived from a different parent

locus

position of a gene on a chromosome

nucleosome

subunit of chromatin composed of a short length of DNA wrapped around a core of histone proteins

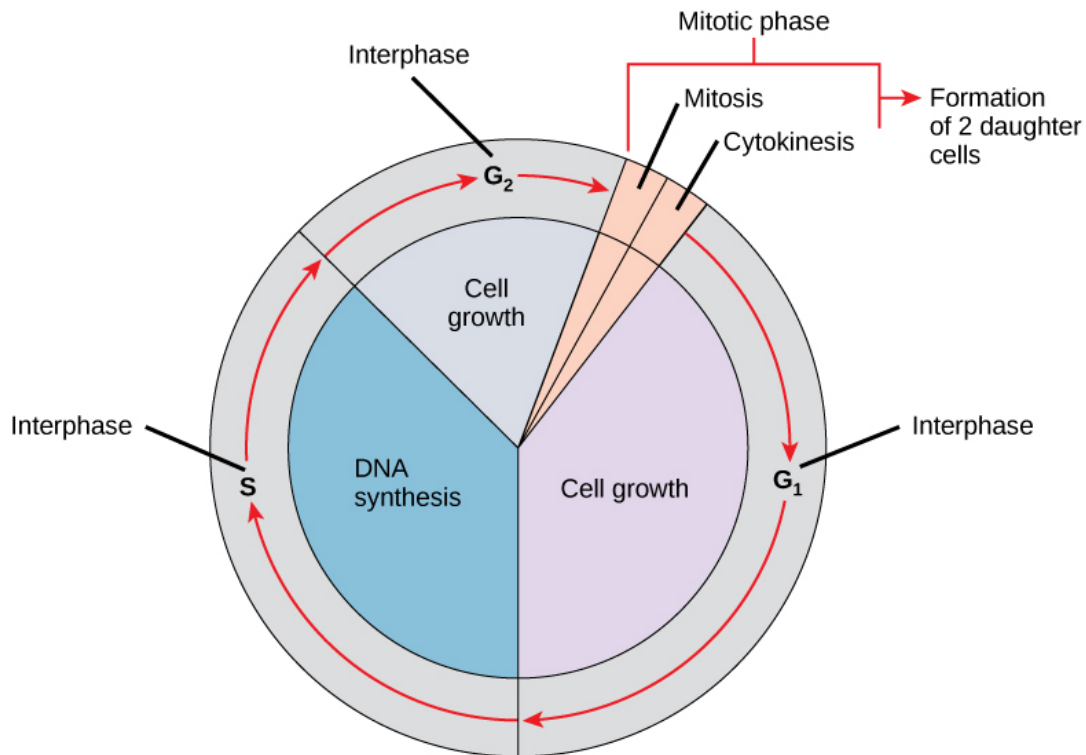


## The Cell Cycle

By the end of this section, you will be able to:

- Describe the three stages of interphase
- Discuss the behavior of chromosomes during karyokinesis
- Explain how the cytoplasmic content is divided during cytokinesis
- Define the quiescent  $G_0$  phase

The **cell cycle** is an ordered series of events involving cell growth and cell division that produces two new daughter cells. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages of growth, DNA replication, and division that produces two identical (clone) cells. The cell cycle has two major phases: interphase and the mitotic phase ([\[link\]](#)). During **interphase**, the cell grows and DNA is replicated. During the **mitotic phase**, the replicated DNA and cytoplasmic contents are separated, and the cell divides.



The cell cycle consists of interphase and the mitotic phase. During interphase, the cell grows and the nuclear DNA is duplicated. Interphase is followed by the mitotic phase. During

the mitotic phase, the duplicated chromosomes are segregated and distributed into daughter nuclei. The cytoplasm is usually divided as well, resulting in two daughter cells.

## Interphase

During interphase, the cell undergoes normal growth processes while also preparing for cell division. In order for a cell to move from interphase into the mitotic phase, many internal and external conditions must be met. The three stages of interphase are called  $G_1$ , S, and  $G_2$ .

### $G_1$ Phase (First Gap)

The first stage of interphase is called the  **$G_1$  phase** (first gap) because, from a microscopic aspect, little change is visible. However, during the  $G_1$  stage, the cell is quite active at the biochemical level. The cell is accumulating the building blocks of chromosomal DNA and the associated proteins as well as accumulating sufficient energy reserves to complete the task of replicating each chromosome in the nucleus.

### S Phase (Synthesis of DNA)

Throughout interphase, nuclear DNA remains in a semi-condensed chromatin configuration. In the **S phase**, DNA replication can proceed through the mechanisms that result in the formation of identical pairs of DNA molecules—sister chromatids—that are firmly attached to the centromeric region. The centrosome is duplicated during the S phase. The two centrosomes will give rise to the **mitotic spindle**, the apparatus that orchestrates the movement of chromosomes during mitosis. At the center of each animal cell, the centrosomes of animal cells are associated with a pair of rod-like objects, the **centrioles**, which are at right angles to each other.

Centrioles help organize cell division. Centrioles are not present in the centrosomes of other eukaryotic species, such as plants and most fungi.

## G<sub>2</sub> Phase (Second Gap)

In the **G<sub>2</sub> phase**, the cell replenishes its energy stores and synthesizes proteins necessary for chromosome manipulation. Some cell organelles are duplicated, and the cytoskeleton is dismantled to provide resources for the mitotic phase. There may be additional cell growth during G<sub>2</sub>. The final preparations for the mitotic phase must be completed before the cell is able to enter the first stage of mitosis.

## The Mitotic Phase

The mitotic phase is a multistep process during which the duplicated chromosomes are aligned, separated, and move into two new, identical daughter cells. The first portion of the mitotic phase is called **karyokinesis**, or nuclear division. The second portion of the mitotic phase, called cytokinesis, is the physical separation of the cytoplasmic components into the two daughter cells.

### Note:

Link to Learning

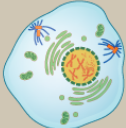
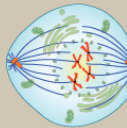
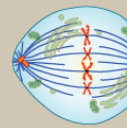
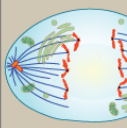
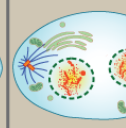
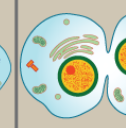
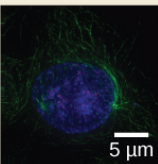
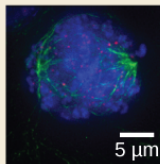
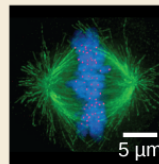
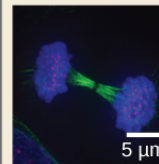
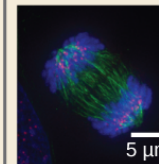
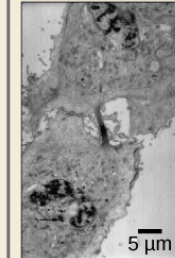


Revisit the stages of mitosis at this [site](#).

# Karyokinesis (Mitosis)

Karyokinesis, also known as **mitosis**, is divided into a series of phases—prophase, prometaphase, metaphase, anaphase, and telophase—that result in the division of the cell nucleus ([link](#)). Karyokinesis is also called mitosis.

**Note:**  
**Art Connection**

Prophase	Prometaphase	Metaphase	Anaphase	Telophase	Cytokinesis
					
<ul style="list-style-type: none"><li>• Chromosomes condense and become visible</li><li>• Spindle fibers emerge from the centrosomes</li><li>• Nuclear envelope breaks down</li><li>• Nucleolus disappears</li></ul>	<ul style="list-style-type: none"><li>• Chromosomes continue to condense</li><li>• Kinetochores appear at the centromeres</li><li>• Mitotic spindle microtubules attach to kinetochores</li><li>• Centrosomes move toward opposite poles</li></ul>	<ul style="list-style-type: none"><li>• Mitotic spindle is fully developed, centrosomes are at opposite poles of the cell</li><li>• Chromosomes are lined up at the metaphase plate</li><li>• Each sister chromatid is attached to a spindle fiber originating from opposite poles</li></ul>	<ul style="list-style-type: none"><li>• Cohesin proteins binding the sister chromatids together break down</li><li>• Sister chromatids (now called chromosomes) are pulled toward opposite poles</li><li>• Non-kinetochore spindle fibers lengthen, elongating the cell</li></ul>	<ul style="list-style-type: none"><li>• Chromosomes arrive at opposite poles and begin to decondense</li><li>• Nuclear envelope material surrounds each set of chromosomes</li><li>• The mitotic spindle breaks down</li></ul>	<ul style="list-style-type: none"><li>• Animal cells: a cleavage furrow separates the daughter cells</li><li>• Plant cells: a cell plate separates the daughter cells</li></ul>
					

MITOSIS

Karyokinesis (or mitosis) is divided into five stages—prophase, prometaphase, metaphase, anaphase, and telophase. The pictures at the bottom were taken by fluorescence microscopy (hence, the black background) of cells artificially stained by fluorescent dyes: blue fluorescence indicates DNA (chromosomes) and green fluorescence indicates microtubules (spindle apparatus). (credit “mitosis drawings”: modification of work by Mariana Ruiz Villareal; credit “micrographs”:

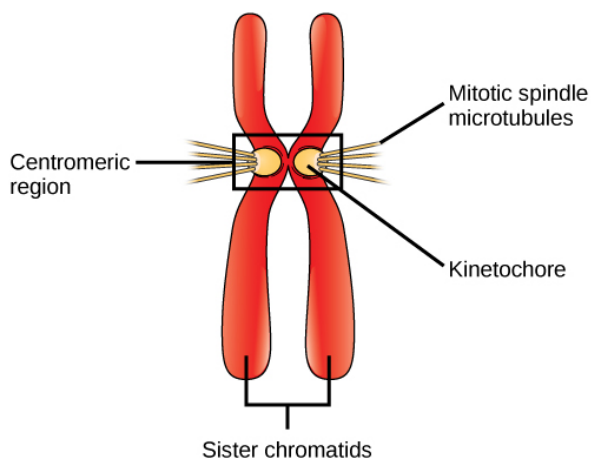
modification of work by Roy van Heesbeen; credit “cytokinesis micrograph”: Wadsworth Center/New York State Department of Health; scale-bar data from Matt Russell)

Which of the following is the correct order of events in mitosis?

- a. Sister chromatids line up at the metaphase plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.
- b. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
- c. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
- d. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

During **prophase**, the “first phase,” the nuclear envelope starts to dissociate into small vesicles, and the membranous organelles (such as the Golgi complex or Golgi apparatus, and endoplasmic reticulum), fragment and disperse toward the periphery of the cell. The nucleolus disappears (disperses). The centrosomes begin to move to opposite poles of the cell. Microtubules that will form the mitotic spindle extend between the centrosomes, pushing them farther apart as the microtubule fibers lengthen. The sister chromatids begin to coil more tightly with the aid of **condensin** proteins and become visible under a light microscope.

During **prometaphase**, the “first change phase,” many processes that were begun in prophase continue to advance. The remnants of the nuclear envelope fragment. The mitotic spindle continues to develop as more microtubules assemble and stretch across the length of the former nuclear area. Chromosomes become more condensed and discrete. Each sister chromatid develops a protein structure called a **kinetochore** in the centromeric region ([\[link\]](#)). The proteins of the kinetochore attract and bind mitotic spindle microtubules. As the spindle microtubules extend from the centrosomes, some of these microtubules come into contact with and firmly bind to the kinetochores. Once a mitotic fiber attaches to a chromosome, the chromosome will be oriented until the kinetochores of sister chromatids face the opposite poles. Eventually, all the sister chromatids will be attached via their kinetochores to microtubules from opposing poles. Spindle microtubules that do not engage the chromosomes are called polar microtubules. These microtubules overlap each other midway between the two poles and contribute to cell elongation. Astral microtubules are located near the poles, aid in spindle orientation, and are required for the regulation of mitosis.



During prometaphase, mitotic spindle microtubules from opposite poles attach to each sister chromatid at the kinetochore. In anaphase, the connection between

the sister chromatids breaks down,  
and the microtubules pull the  
chromosomes toward opposite  
poles.

During **metaphase**, the “change phase,” all the chromosomes are aligned in a plane called the **metaphase plate**, or the equatorial plane, midway between the two poles of the cell. The sister chromatids are still tightly attached to each other by cohesin proteins. At this time, the chromosomes are maximally condensed.

During **anaphase**, the “upward phase,” the cohesin proteins degrade, and the sister chromatids separate at the centromere. Each chromatid, now called a chromosome, is pulled rapidly toward the centrosome to which its microtubule is attached. The cell becomes visibly elongated (oval shaped) as the polar microtubules slide against each other at the metaphase plate where they overlap.

During **telophase**, the “distance phase,” the chromosomes reach the opposite poles and begin to decondense (unravel), relaxing into a chromatin configuration. The mitotic spindles are depolymerized into tubulin monomers that will be used to assemble cytoskeletal components for each daughter cell. Nuclear envelopes form around the chromosomes, and nucleosomes appear within the nuclear area.

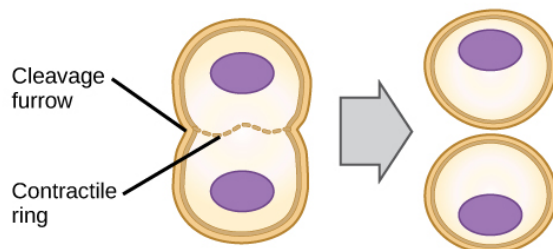
## **Cytokinesis**

**Cytokinesis**, or “cell motion,” is the second main stage of the mitotic phase during which cell division is completed via the physical separation of the cytoplasmic components into two daughter cells. Division is not complete until the cell components have been apportioned and completely separated into the two daughter cells. Although the stages of mitosis are similar for most eukaryotes, the process of cytokinesis is quite different for eukaryotes that have cell walls, such as plant cells.

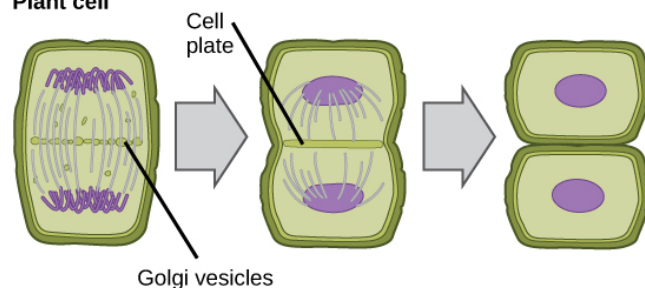
In cells such as animal cells that lack cell walls, cytokinesis follows the onset of anaphase. A contractile ring composed of actin filaments forms just inside the plasma membrane at the former metaphase plate. The actin filaments pull the equator of the cell inward, forming a fissure. This fissure, or “crack,” is called the **cleavage furrow**. The furrow deepens as the actin ring contracts, and eventually the membrane is cleaved in two ([\[link\]](#)).

In plant cells, a new cell wall must form between the daughter cells. During interphase, the Golgi apparatus accumulates enzymes, structural proteins, and glucose molecules prior to breaking into vesicles and dispersing throughout the dividing cell. During telophase, these Golgi vesicles are transported on microtubules to form a phragmoplast (a vesicular structure) at the metaphase plate. There, the vesicles fuse and coalesce from the center toward the cell walls; this structure is called a **cell plate**. As more vesicles fuse, the cell plate enlarges until it merges with the cell walls at the periphery of the cell. Enzymes use the glucose that has accumulated between the membrane layers to build a new cell wall. The Golgi membranes become parts of the plasma membrane on either side of the new cell wall ([\[link\]](#)).

**Animal cell**



**Plant cell**



During cytokinesis in animal cells,



a ring of actin filaments forms at the metaphase plate. The ring contracts, forming a cleavage furrow, which divides the cell in two. In plant cells, Golgi vesicles coalesce at the former metaphase plate, forming a phragmoplast. A cell plate formed by the fusion of the vesicles of the phragmoplast grows from the center toward the cell walls, and the membranes of the vesicles fuse to form a plasma membrane that divides the cell in two.

## **G<sub>0</sub> Phase**

Not all cells adhere to the classic cell cycle pattern in which a newly formed daughter cell immediately enters the preparatory phases of interphase, closely followed by the mitotic phase. Cells in **G<sub>0</sub> phase** are not actively preparing to divide. The cell is in a **quiescent** (inactive) stage that occurs when cells exit the cell cycle. Some cells enter G<sub>0</sub> temporarily until an external signal triggers the onset of G<sub>1</sub>. Other cells that never or rarely divide, such as mature cardiac muscle and nerve cells, remain in G<sub>0</sub> permanently.

### **Note:**

Scientific Method Connection

### **Determine the Time Spent in Cell Cycle Stages**

**Problem:** How long does a cell spend in interphase compared to each stage of mitosis?

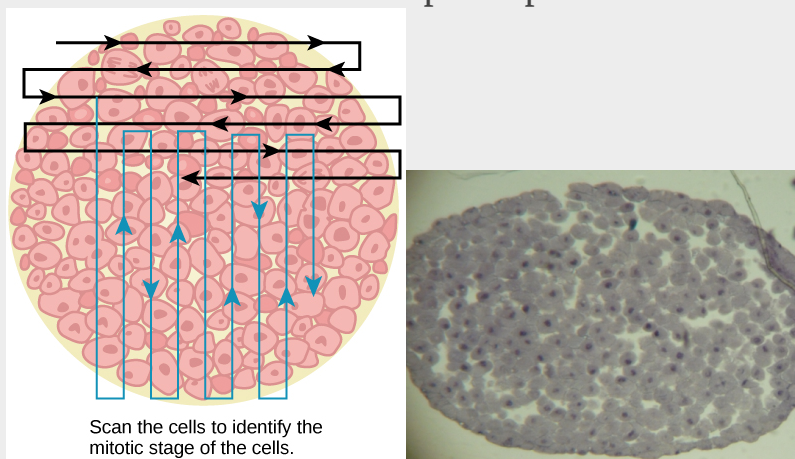
**Background:** A prepared microscope slide of blastula cross-sections will show cells arrested in various stages of the cell cycle. It is not visually

possible to separate the stages of interphase from each other, but the mitotic stages are readily identifiable. If 100 cells are examined, the number of cells in each identifiable cell cycle stage will give an estimate of the time it takes for the cell to complete that stage.

**Problem Statement:** Given the events included in all of interphase and those that take place in each stage of mitosis, estimate the length of each stage based on a 24-hour cell cycle. Before proceeding, state your hypothesis.

**Test your hypothesis:** Test your hypothesis by doing the following:

1. Place a fixed and stained microscope slide of whitefish blastula cross-sections under the scanning objective of a light microscope.
2. Locate and focus on one of the sections using the scanning objective of your microscope. Notice that the section is a circle composed of dozens of closely packed individual cells.
3. Switch to the low-power objective and refocus. With this objective, individual cells are visible.
4. Switch to the high-power objective and slowly move the slide left to right, and up and down to view all the cells in the section ([\[link\]](#)). As you scan, you will notice that most of the cells are not undergoing mitosis but are in the interphase period of the cell cycle.



Slowly scan whitefish blastula cells with the high-power objective as illustrated in image (a) to identify their mitotic stage. (b) A microscopic image of the scanned cells is

shown. (credit “micrograph”: modification of work by Linda Flora; scale-bar data from Matt Russell)

5. Practice identifying the various stages of the cell cycle, using the drawings of the stages as a guide ([\[link\]](#)).
6. Once you are confident about your identification, begin to record the stage of each cell you encounter as you scan left to right, and top to bottom across the blastula section.
7. Keep a tally of your observations and stop when you reach 100 cells identified.
8. The larger the sample size (total number of cells counted), the more accurate the results. If possible, gather and record group data prior to calculating percentages and making estimates.

**Record your observations:** Make a table similar to [\[link\]](#) in which you record your observations.

### Results of Cell Stage Identification

Phase or Stage	Individual Totals	Group Totals	Percent
Interphase			
Prophase			
Metaphase			
Anaphase			

Results of Cell Stage Identification			
Phase or Stage	Individual Totals	Group Totals	Percent
Telophase			
Cytokinesis			
Totals	100	100	100 percent

Analyze your data/report your results: To find the length of time whitefish blastula cells spend in each stage, multiply the percent (recorded as a decimal) by 24 hours. Make a table similar to [\[link\]](#) to illustrate your data.

Estimate of Cell Stage Length		
Phase or Stage	Percent (as Decimal)	Time in Hours
Interphase		
Prophase		
Metaphase		
Anaphase		
Telophase		

Estimate of Cell Stage Length		
Phase or Stage	Percent (as Decimal)	Time in Hours
Cytokinesis		
<b>Draw a conclusion:</b> Did your results support your estimated times? Were any of the outcomes unexpected? If so, discuss which events in that stage might contribute to the calculated time.		

## Section Summary

The cell cycle is an orderly sequence of events. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages. In eukaryotes, the cell cycle consists of a long preparatory period, called interphase. Interphase is divided into  $G_1$ , S, and  $G_2$  phases. The mitotic phase begins with karyokinesis (mitosis), which consists of five stages: prophase, prometaphase, metaphase, anaphase, and telophase. The final stage of the mitotic phase is cytokinesis, during which the cytoplasmic components of the daughter cells are separated either by an actin ring (animal cells) or by cell plate formation (plant cells).

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Which of the following is the correct order of events in mitosis?

- Sister chromatids line up at the metaphase plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.

- b. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
- c. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
- d. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

---

**Solution:**

[\[link\]](#) D. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

## Review Questions

**Exercise:**

**Problem:**

Chromosomes are duplicated during what stage of the cell cycle?

- a. G<sub>1</sub> phase
- b. S phase
- c. prophase
- d. prometaphase

---

**Solution:**

B

**Exercise:**

**Problem:**

Which of the following events does not occur during some stages of interphase?

- a. DNA duplication
- b. organelle duplication
- c. increase in cell size
- d. separation of sister chromatids

---

**Solution:**

D

**Exercise:**

**Problem:** The mitotic spindles arise from which cell structure?

- a. centromere
- b. centrosome
- c. kinetochore
- d. cleavage furrow

---

**Solution:**

B

**Exercise:**

**Problem:**

Attachment of the mitotic spindle fibers to the kinetochores is a characteristic of which stage of mitosis?

- a. prophase
- b. prometaphase

- c. metaphase
- d. anaphase

---

**Solution:**

B

**Exercise:**

**Problem:**

Unpacking of chromosomes and the formation of a new nuclear envelope is a characteristic of which stage of mitosis?

- a. prometaphase
- b. metaphase
- c. anaphase
- d. telophase

---

**Solution:**

D

**Exercise:**

**Problem:**

Separation of the sister chromatids is a characteristic of which stage of mitosis?

- a. prometaphase
- b. metaphase
- c. anaphase
- d. telophase

---

**Solution:**

C



**Exercise:**

**Problem:**

The chromosomes become visible under a light microscope during which stage of mitosis?

- a. prophase
- b. prometaphase
- c. metaphase
- d. anaphase

---

**Solution:**

A

**Exercise:**

**Problem:**

The fusing of Golgi vesicles at the metaphase plate of dividing plant cells forms what structure?

- a. cell plate
- b. actin ring
- c. cleavage furrow
- d. mitotic spindle

---

**Solution:**

A

**Free Response**

**Exercise:**

**Problem:**

Briefly describe the events that occur in each phase of interphase.

---

**Solution:**

During  $G_1$ , the cell increases in size, the genomic DNA is assessed for damage, and the cell stockpiles energy reserves and the components to synthesize DNA. During the S phase, the chromosomes, the centrosomes, and the centrioles (animal cells) duplicate. During the  $G_2$  phase, the cell recovers from the S phase, continues to grow, duplicates some organelles, and dismantles other organelles.

**Exercise:****Problem:**

Chemotherapy drugs such as vincristine and colchicine disrupt mitosis by binding to tubulin (the subunit of microtubules) and interfering with microtubule assembly and disassembly. Exactly what mitotic structure is targeted by these drugs and what effect would that have on cell division?

---

**Solution:**

The mitotic spindle is formed of microtubules. Microtubules are polymers of the protein tubulin; therefore, it is the mitotic spindle that is disrupted by these drugs. Without a functional mitotic spindle, the chromosomes will not be sorted or separated during mitosis. The cell will arrest in mitosis and die.

**Exercise:****Problem:**

Describe the similarities and differences between the cytokinesis mechanisms found in animal cells versus those in plant cells.

---

**Solution:**

There are very few similarities between animal cell and plant cell cytokinesis. In animal cells, a ring of actin fibers is formed around the periphery of the cell at the former metaphase plate (cleavage furrow). The actin ring contracts inward, pulling the plasma membrane toward the center of the cell until the cell is pinched in two. In plant cells, a new cell wall must be formed between the daughter cells. Due to the rigid cell walls of the parent cell, contraction of the middle of the cell is not possible. Instead, a phragmoplast first forms. Subsequently, a cell plate is formed in the center of the cell at the former metaphase plate. The cell plate is formed from Golgi vesicles that contain enzymes, proteins, and glucose. The vesicles fuse and the enzymes build a new cell wall from the proteins and glucose. The cell plate grows toward and eventually fuses with the cell wall of the parent cell.

**Exercise:**

**Problem:**

List some reasons why a cell that has just completed cytokinesis might enter the  $G_0$  phase instead of the  $G_1$  phase.

---

**Solution:**

Many cells temporarily enter  $G_0$  until they reach maturity. Some cells are only triggered to enter  $G_1$  when the organism needs to increase that particular cell type. Some cells only reproduce following an injury to the tissue. Some cells never divide once they reach maturity.

**Exercise:**

**Problem:**

What cell cycle events will be affected in a cell that produces mutated (non-functional) cohesin protein?

---

**Solution:**

If cohesin is not functional, chromosomes are not packaged after DNA replication in the S phase of interphase. It is likely that the proteins of the centromeric region, such as the kinetochore, would not form. Even

if the mitotic spindle fibers could attach to the chromatids without packing, the chromosomes would not be sorted or separated during mitosis.

## **Glossary**

### **anaphase**

stage of mitosis during which sister chromatids are separated from each other

### **cell cycle**

ordered series of events involving cell growth and cell division that produces two new daughter cells

### **cell plate**

structure formed during plant cell cytokinesis by Golgi vesicles, forming a temporary structure (phragmoplast) and fusing at the metaphase plate; ultimately leads to the formation of cell walls that separate the two daughter cells

### **centriole**

rod-like structure constructed of microtubules at the center of each animal cell centrosome

### **cleavage furrow**

constriction formed by an actin ring during cytokinesis in animal cells that leads to cytoplasmic division

### **condensin**

proteins that help sister chromatids coil during prophase

### **cytokinesis**

division of the cytoplasm following mitosis that forms two daughter cells.

### **G<sub>0</sub> phase**

distinct from the  $G_1$  phase of interphase; a cell in  $G_0$  is not preparing to divide

$G_1$  phase

(also, first gap) first phase of interphase centered on cell growth during mitosis

$G_2$  phase

(also, second gap) third phase of interphase during which the cell undergoes final preparations for mitosis

interphase

period of the cell cycle leading up to mitosis; includes  $G_1$ , S, and  $G_2$  phases (the interim period between two consecutive cell divisions)

karyokinesis

mitotic nuclear division

kinetochore

protein structure associated with the centromere of each sister chromatid that attracts and binds spindle microtubules during prometaphase

metaphase plate

equatorial plane midway between the two poles of a cell where the chromosomes align during metaphase

metaphase

stage of mitosis during which chromosomes are aligned at the metaphase plate

mitosis

(also, karyokinesis) period of the cell cycle during which the duplicated chromosomes are separated into identical nuclei; includes prophase, prometaphase, metaphase, anaphase, and telophase

mitotic phase

period of the cell cycle during which duplicated chromosomes are distributed into two nuclei and cytoplasmic contents are divided; includes karyokinesis (mitosis) and cytokinesis

mitotic spindle

apparatus composed of microtubules that orchestrates the movement of chromosomes during mitosis

prometaphase

stage of mitosis during which the nuclear membrane breaks down and mitotic spindle fibers attach to kinetochores

prophase

stage of mitosis during which chromosomes condense and the mitotic spindle begins to form

quiescent

refers to a cell that is performing normal cell functions and has not initiated preparations for cell division

S phase

second, or synthesis, stage of interphase during which DNA replication occurs

telophase

stage of mitosis during which chromosomes arrive at opposite poles, decondense, and are surrounded by a new nuclear envelope

## Control of the Cell Cycle

By the end of this section, you will be able to:

- Understand how the cell cycle is controlled by mechanisms both internal and external to the cell
- Explain how the three internal control checkpoints occur at the end of  $G_1$ , at the  $G_2/M$  transition, and during metaphase
- Describe the molecules that control the cell cycle through positive and negative regulation

The length of the cell cycle is highly variable, even within the cells of a single organism. In humans, the frequency of cell turnover ranges from a few hours in early embryonic development, to an average of two to five days for epithelial cells, and to an entire human lifetime spent in  $G_0$  by specialized cells, such as cortical neurons or cardiac muscle cells. There is also variation in the time that a cell spends in each phase of the cell cycle. When fast-dividing mammalian cells are grown in culture (outside the body under optimal growing conditions), the length of the cycle is about 24 hours. In rapidly dividing human cells with a 24-hour cell cycle, the  $G_1$  phase lasts approximately nine hours, the S phase lasts 10 hours, the  $G_2$  phase lasts about four and one-half hours, and the M phase lasts approximately one-half hour. In early embryos of fruit flies, the cell cycle is completed in about eight minutes. The timing of events in the cell cycle is controlled by mechanisms that are both internal and external to the cell.

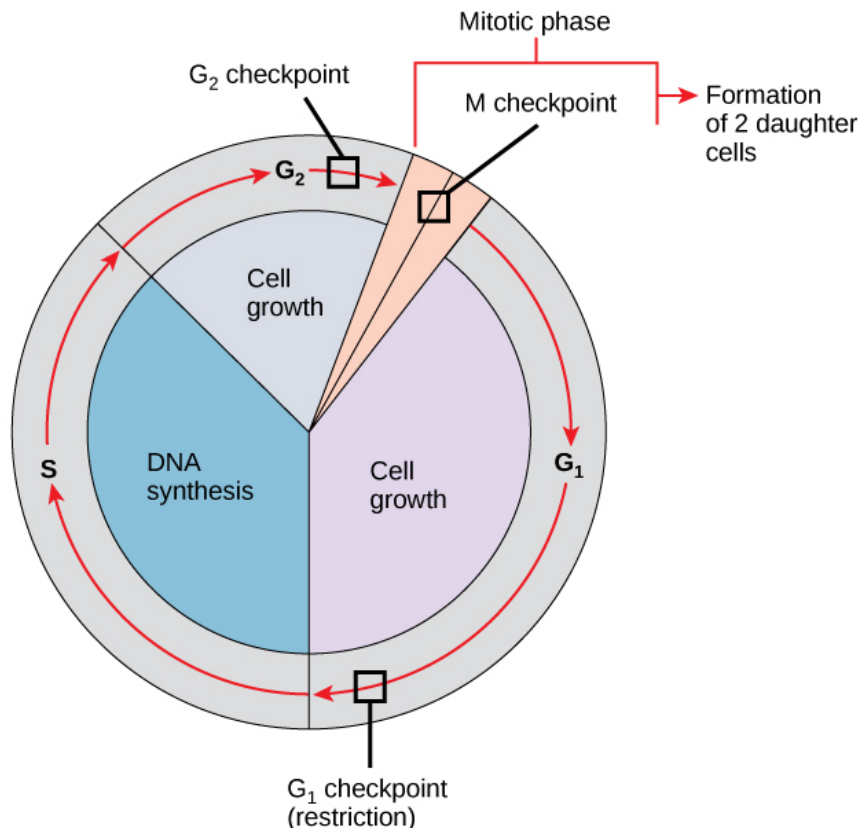
## Regulation of the Cell Cycle by External Events

Both the initiation and inhibition of cell division are triggered by events external to the cell when it is about to begin the replication process. An event may be as simple as the death of a nearby cell or as sweeping as the release of growth-promoting hormones, such as human growth hormone (HGH). A lack of HGH can inhibit cell division, resulting in dwarfism, whereas too much HGH can result in gigantism. Crowding of cells can also inhibit cell division. Another factor that can initiate cell division is the size of the cell; as a cell grows, it becomes inefficient due to its decreasing surface-to-volume ratio. The solution to this problem is to divide.

Whatever the source of the message, the cell receives the signal, and a series of events within the cell allows it to proceed into interphase. Moving forward from this initiation point, every parameter required during each cell cycle phase must be met or the cycle cannot progress.

## Regulation at Internal Checkpoints

It is essential that the daughter cells produced be exact duplicates of the parent cell. Mistakes in the duplication or distribution of the chromosomes lead to mutations that may be passed forward to every new cell produced from an abnormal cell. To prevent a compromised cell from continuing to divide, there are internal control mechanisms that operate at three main **cell cycle checkpoints**. A checkpoint is one of several points in the eukaryotic cell cycle at which the progression of a cell to the next stage in the cycle can be halted until conditions are favorable. These checkpoints occur near the end of  $G_1$ , at the  $G_2/M$  transition, and during metaphase ([\[link\]](#)).





The cell cycle is controlled at three checkpoints. The integrity of the DNA is assessed at the  $G_1$  checkpoint. Proper chromosome duplication is assessed at the  $G_2$  checkpoint. Attachment of each kinetochore to a spindle fiber is assessed at the M checkpoint.

### **The $G_1$ Checkpoint**

The  $G_1$  checkpoint determines whether all conditions are favorable for cell division to proceed. The  $G_1$  checkpoint, also called the restriction point (in yeast), is a point at which the cell irreversibly commits to the cell division process. External influences, such as growth factors, play a large role in carrying the cell past the  $G_1$  checkpoint. In addition to adequate reserves and cell size, there is a check for genomic DNA damage at the  $G_1$  checkpoint. A cell that does not meet all the requirements will not be allowed to progress into the S phase. The cell can halt the cycle and attempt to remedy the problematic condition, or the cell can advance into  $G_0$  and await further signals when conditions improve.

### **The $G_2$ Checkpoint**

The  $G_2$  checkpoint bars entry into the mitotic phase if certain conditions are not met. As at the  $G_1$  checkpoint, cell size and protein reserves are assessed. However, the most important role of the  $G_2$  checkpoint is to ensure that all of the chromosomes have been replicated and that the replicated DNA is not damaged. If the checkpoint mechanisms detect problems with the DNA, the cell cycle is halted, and the cell attempts to either complete DNA replication or repair the damaged DNA.

### **The M Checkpoint**

The M checkpoint occurs near the end of the metaphase stage of karyokinesis. The M checkpoint is also known as the spindle checkpoint, because it determines whether all the sister chromatids are correctly attached to the spindle microtubules. Because the separation of the sister chromatids during anaphase is an irreversible step, the cycle will not proceed until the kinetochores of each pair of sister chromatids are firmly anchored to at least two spindle fibers arising from opposite poles of the cell.

**Note:**

Link to Learning



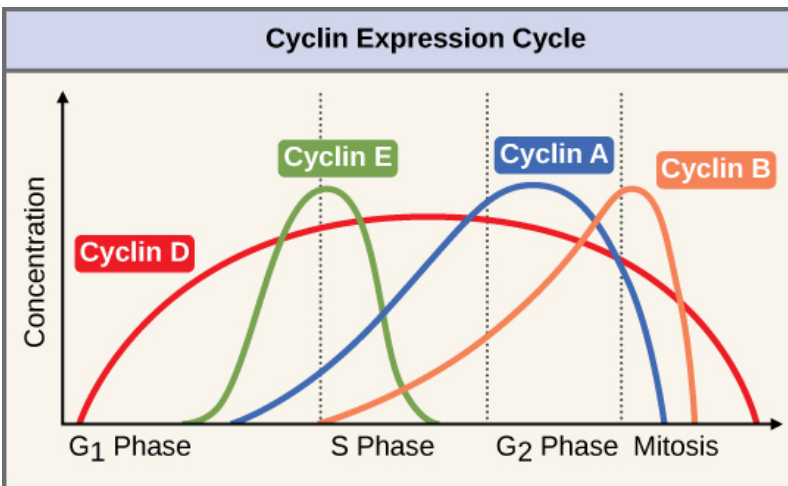
Watch what occurs at the  $G_1$ ,  $G_2$ , and M checkpoints by visiting this [website](#) to see an animation of the cell cycle.

## Regulator Molecules of the Cell Cycle

In addition to the internally controlled checkpoints, there are two groups of intracellular molecules that regulate the cell cycle. These regulatory molecules either promote progress of the cell to the next phase (positive regulation) or halt the cycle (negative regulation). Regulator molecules may act individually, or they can influence the activity or production of other regulatory proteins. Therefore, the failure of a single regulator may have almost no effect on the cell cycle, especially if more than one mechanism controls the same event. Conversely, the effect of a deficient or non-functioning regulator can be wide-ranging and possibly fatal to the cell if multiple processes are affected.

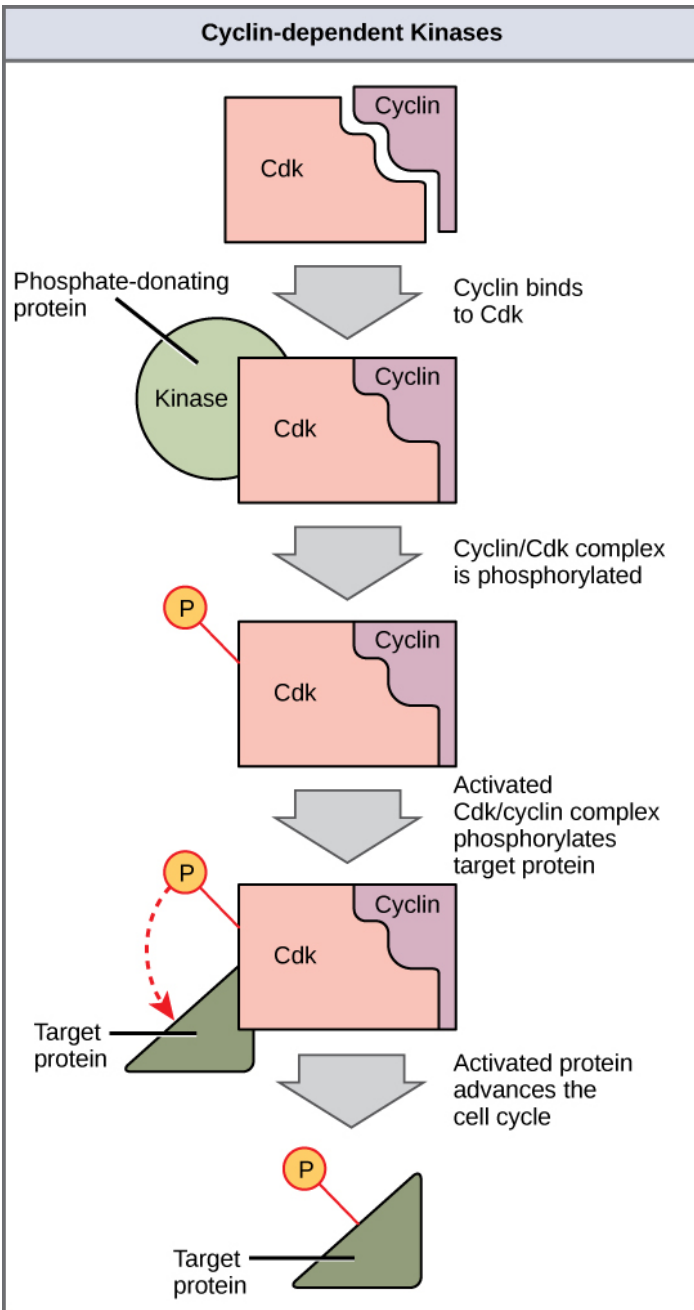
## Positive Regulation of the Cell Cycle

Two groups of proteins, called **cyclins** and **cyclin-dependent kinases** (Cdks), are responsible for the progress of the cell through the various checkpoints. The levels of the four cyclin proteins fluctuate throughout the cell cycle in a predictable pattern ([\[link\]](#)). Increases in the concentration of cyclin proteins are triggered by both external and internal signals. After the cell moves to the next stage of the cell cycle, the cyclins that were active in the previous stage are degraded.



The concentrations of cyclin proteins change throughout the cell cycle. There is a direct correlation between cyclin accumulation and the three major cell cycle checkpoints. Also note the sharp decline of cyclin levels following each checkpoint (the transition between phases of the cell cycle), as cyclin is degraded by cytoplasmic enzymes.  
(credit: modification of work by "WikiMiMa"/Wikimedia Commons)

Cyclins regulate the cell cycle only when they are tightly bound to Cdks. To be fully active, the Cdk/cyclin complex must also be phosphorylated in specific locations. Like all kinases, Cdks are enzymes (kinases) that phosphorylate other proteins. Phosphorylation activates the protein by changing its shape. The proteins phosphorylated by Cdks are involved in advancing the cell to the next phase. ([link](#)). The levels of Cdk proteins are relatively stable throughout the cell cycle; however, the concentrations of cyclin fluctuate and determine when Cdk/cyclin complexes form. The different cyclins and Cdks bind at specific points in the cell cycle and thus regulate different checkpoints.



Cyclin-dependent kinases (Cdks) are protein kinases that, when fully activated, can phosphorylate and thus activate other proteins that advance the cell cycle past a checkpoint. To become fully activated, a Cdk must bind to a cyclin protein and then be phosphorylated by another kinase.

Since the cyclic fluctuations of cyclin levels are based on the timing of the cell cycle and not on specific events, regulation of the cell cycle usually occurs by either the Cdk molecules alone or the Cdk/cyclin complexes. Without a specific concentration of fully activated cyclin/Cdk complexes, the cell cycle cannot proceed through the checkpoints.

Although the cyclins are the main regulatory molecules that determine the forward momentum of the cell cycle, there are several other mechanisms that fine-tune the progress of the cycle with negative, rather than positive, effects. These mechanisms essentially block the progression of the cell cycle until problematic conditions are resolved. Molecules that prevent the full activation of Cdks are called Cdk inhibitors. Many of these inhibitor molecules directly or indirectly monitor a particular cell cycle event. The block placed on Cdks by inhibitor molecules will not be removed until the specific event that the inhibitor monitors is completed.

## **Negative Regulation of the Cell Cycle**

The second group of cell cycle regulatory molecules are negative regulators. Negative regulators halt the cell cycle. Remember that in positive regulation, active molecules cause the cycle to progress.

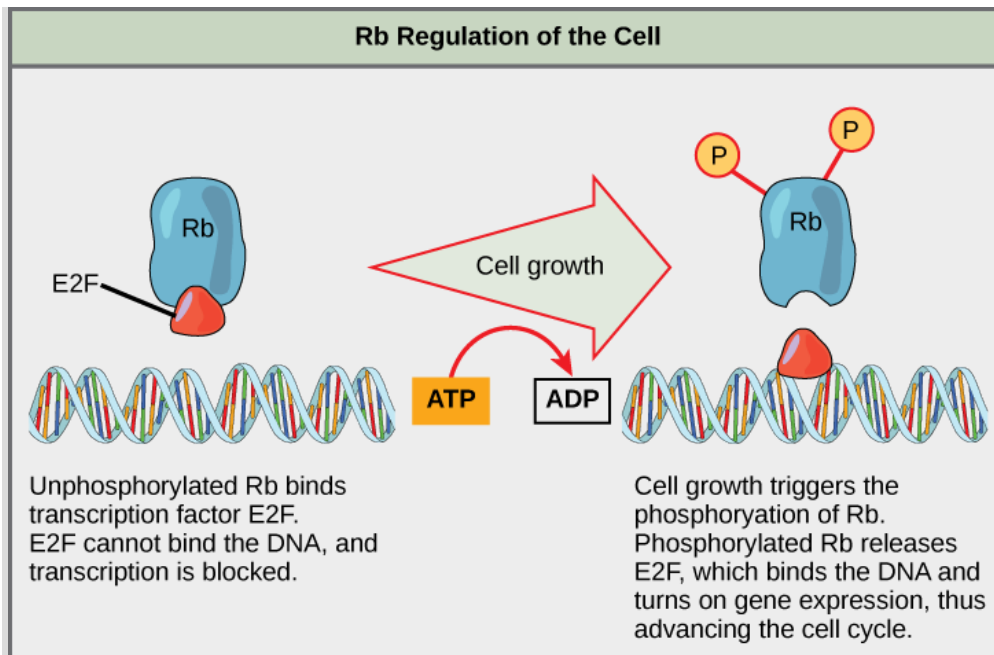
The best understood negative regulatory molecules are **retinoblastoma protein (Rb)**, **p53**, and **p21**. Retinoblastoma proteins are a group of tumor-suppressor proteins common in many cells. The 53 and 21 designations refer to the functional molecular masses of the proteins (p) in kilodaltons. Much of what is known about cell cycle regulation comes from research conducted with cells that have lost regulatory control. All three of these regulatory proteins were discovered to be damaged or non-functional in cells that had begun to replicate uncontrollably (became cancerous). In each case, the main cause of the unchecked progress through the cell cycle was a faulty copy of the regulatory protein.

Rb, p53, and p21 act primarily at the G<sub>1</sub> checkpoint. p53 is a multi-functional protein that has a major impact on the commitment of a cell to division because it acts when there is damaged DNA in cells that are undergoing the preparatory processes during G<sub>1</sub>. If damaged DNA is detected, p53 halts the cell cycle and recruits enzymes to repair the DNA. If the DNA cannot be repaired, p53 can trigger apoptosis, or cell suicide, to prevent the duplication of damaged chromosomes. As p53 levels rise, the production of p21 is triggered. p21 enforces the halt in the cycle dictated by p53 by binding to and inhibiting the activity of the Cdk/cyclin complexes. As a cell is exposed to more stress, higher levels of p53 and p21 accumulate, making it less likely that the cell will move into the S phase.

Rb exerts its regulatory influence on other positive regulator proteins. Chiefly, Rb monitors cell size. In the active, dephosphorylated state, Rb binds to proteins called transcription factors, most commonly, E2F ([link](#)). Transcription factors “turn on” specific genes, allowing the production of proteins encoded by that gene. When Rb is bound to E2F, production of proteins necessary for the G<sub>1</sub>/S transition is blocked. As the cell increases in size, Rb is slowly phosphorylated until it becomes inactivated. Rb releases E2F, which can now turn on the gene that produces the transition protein, and this particular block is removed. For the cell to move past each of the checkpoints, all positive regulators must be “turned on,” and all negative regulators must be “turned off.”

**Note:**

Art Connection



Rb halts the cell cycle and releases its hold in response to cell growth.

Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be appropriate for these proteins?

## Section Summary

Each step of the cell cycle is monitored by internal controls called checkpoints. There are three major checkpoints in the cell cycle: one near the end of  $G_1$ , a second at the  $G_2/M$  transition, and the third during metaphase. Positive regulator molecules allow the cell cycle to advance to the next stage. Negative regulator molecules monitor cellular conditions and can halt the cycle until specific requirements are met.

## Art Connections



**Exercise:****Problem:**

[\[link\]](#) Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be an appropriate for these proteins?

---

**Solution:**

[\[link\]](#) Rb and other negative regulatory proteins control cell division and therefore prevent the formation of tumors. Mutations that prevent these proteins from carrying out their function can result in cancer.

**Review Questions****Exercise:****Problem:**

At which of the cell cycle checkpoints do external forces have the greatest influence?

- a. G<sub>1</sub> checkpoint
  - b. G<sub>2</sub> checkpoint
  - c. M checkpoint
  - d. G<sub>0</sub> checkpoint
- 

**Solution:**

A

**Exercise:****Problem:**

What is the main prerequisite for clearance at the G<sub>2</sub> checkpoint?

- a. cell has reached a sufficient size

- b. an adequate stockpile of nucleotides
- c. accurate and complete DNA replication
- d. proper attachment of mitotic spindle fibers to kinetochores

---

**Solution:**

C

**Exercise:**

**Problem:**

If the M checkpoint is not cleared, what stage of mitosis will be blocked?

- a. prophase
- b. prometaphase
- c. metaphase
- d. anaphase

---

**Solution:**

D

**Exercise:**

**Problem:**

Which protein is a positive regulator that phosphorylates other proteins when activated?

- a. p53
- b. retinoblastoma protein (Rb)
- c. cyclin
- d. cyclin-dependent kinase (Cdk)

---

**Solution:**

D

**Exercise:**

**Problem:**

Many of the negative regulator proteins of the cell cycle were discovered in what type of cells?

- a. gametes
- b. cells in  $G_0$
- c. cancer cells
- d. stem cells

---

**Solution:**

C

**Exercise:**

**Problem:**

Which negative regulatory molecule can trigger cell suicide (apoptosis) if vital cell cycle events do not occur?

- a. p53
- b. p21
- c. retinoblastoma protein (Rb)
- d. cyclin-dependent kinase (Cdk)

---

**Solution:**

A

**Free Response**

**Exercise:**

**Problem:**

Describe the general conditions that must be met at each of the three main cell cycle checkpoints.

---

**Solution:**

The G<sub>1</sub> checkpoint monitors adequate cell growth, the state of the genomic DNA, adequate stores of energy, and materials for S phase. At the G<sub>2</sub> checkpoint, DNA is checked to ensure that all chromosomes were duplicated and that there are no mistakes in newly synthesized DNA. Additionally, cell size and energy reserves are evaluated. The M checkpoint confirms the correct attachment of the mitotic spindle fibers to the kinetochores.

**Exercise:****Problem:**

Explain the roles of the positive cell cycle regulators compared to the negative regulators.

---

**Solution:**

Positive cell regulators such as cyclin and Cdk perform tasks that advance the cell cycle to the next stage. Negative regulators such as Rb, p53, and p21 block the progression of the cell cycle until certain events have occurred.

**Exercise:**

**Problem:** What steps are necessary for Cdk to become fully active?

---

**Solution:**

Cdk must bind to a cyclin, and it must be phosphorylated in the correct position to become fully active.

**Exercise:**

**Problem:**

Rb is a negative regulator that blocks the cell cycle at the G<sub>1</sub> checkpoint until the cell achieves a requisite size. What molecular mechanism does Rb employ to halt the cell cycle?

---

**Solution:**

Rb is active when it is dephosphorylated. In this state, Rb binds to E2F, which is a transcription factor required for the transcription and eventual translation of molecules required for the G<sub>1</sub>/S transition. E2F cannot transcribe certain genes when it is bound to Rb. As the cell increases in size, Rb becomes phosphorylated, inactivated, and releases E2F. E2F can then promote the transcription of the genes it controls, and the transition proteins will be produced.

**Glossary**

cell cycle checkpoint

mechanism that monitors the preparedness of a eukaryotic cell to advance through the various cell cycle stages

cyclin

one of a group of proteins that act in conjunction with cyclin-dependent kinases to help regulate the cell cycle by phosphorylating key proteins; the concentrations of cyclins fluctuate throughout the cell cycle

cyclin-dependent kinase

one of a group of protein kinases that helps to regulate the cell cycle when bound to cyclin; it functions to phosphorylate other proteins that are either activated or inactivated by phosphorylation

p21

cell cycle regulatory protein that inhibits the cell cycle; its levels are controlled by p53

p53

cell cycle regulatory protein that regulates cell growth and monitors DNA damage; it halts the progression of the cell cycle in cases of DNA damage and may induce apoptosis

retinoblastoma protein (Rb)

regulatory molecule that exhibits negative effects on the cell cycle by interacting with a transcription factor (E2F)

## Prokaryotic Cell Division

By the end of this section, you will be able to:

- Describe the process of binary fission in prokaryotes
- Explain how FtsZ and tubulin proteins are examples of homology

Prokaryotes, such as bacteria, propagate by binary fission. For unicellular organisms, cell division is the only method to produce new individuals. In both prokaryotic and eukaryotic cells, the outcome of cell reproduction is a pair of daughter cells that are genetically identical to the parent cell. In unicellular organisms, daughter cells are individuals.

To achieve the outcome of cloned offspring, certain steps are essential. The genomic DNA must be replicated and then allocated into the daughter cells; the cytoplasmic contents must also be divided to give both new cells the machinery to sustain life. In bacterial cells, the genome consists of a single, circular DNA chromosome; therefore, the process of cell division is simplified. Karyokinesis is unnecessary because there is no nucleus and thus no need to direct one copy of the multiple chromosomes into each daughter cell. This type of cell division is called **binary (prokaryotic) fission**.

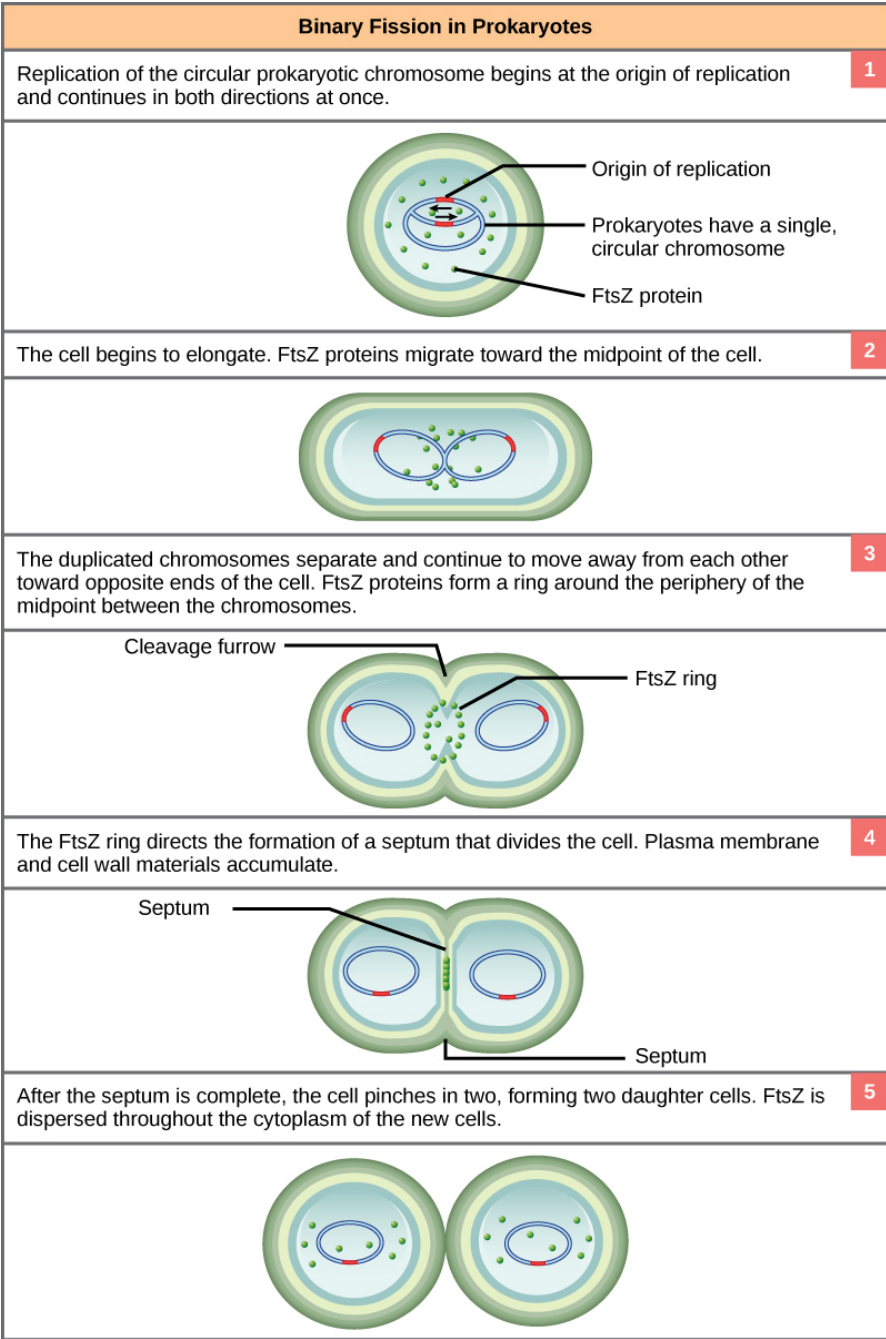
## Binary Fission

Due to the relative simplicity of the prokaryotes, the cell division process, called binary fission, is a less complicated and much more rapid process than cell division in eukaryotes. The single, circular DNA chromosome of bacteria is not enclosed in a nucleus, but instead occupies a specific location, the nucleoid, within the cell ([\[link\]](#)). Although the DNA of the nucleoid is associated with proteins that aid in packaging the molecule into a compact size, there are no histone proteins and thus no nucleosomes in prokaryotes. The packing proteins of bacteria are, however, related to the cohesin and condensin proteins involved in the chromosome compaction of eukaryotes.

The bacterial chromosome is attached to the plasma membrane at about the midpoint of the cell. The starting point of replication, the **origin**, is close to the binding site of the chromosome to the plasma membrane ([\[link\]](#)). Replication of the DNA is bidirectional, moving away from the origin on both strands of the loop simultaneously. As the new double strands are

formed, each origin point moves away from the cell wall attachment toward the opposite ends of the cell. As the cell elongates, the growing membrane aids in the transport of the chromosomes. After the chromosomes have cleared the midpoint of the elongated cell, cytoplasmic separation begins. The formation of a ring composed of repeating units of a protein called **FtsZ** directs the partition between the nucleoids. Formation of the FtsZ ring triggers the accumulation of other proteins that work together to recruit new membrane and cell wall materials to the site. A **septum** is formed between the nucleoids, extending gradually from the periphery toward the center of the cell. When the new cell walls are in place, the daughter cells separate.





These images show the steps of binary fission in prokaryotes. (credit: modification of work by “Mcstrother”/Wikimedia Commons)

**Note:****Evolution Connection****Mitotic Spindle Apparatus**

The precise timing and formation of the mitotic spindle is critical to the success of eukaryotic cell division. Prokaryotic cells, on the other hand, do not undergo karyokinesis and therefore have no need for a mitotic spindle. However, the FtsZ protein that plays such a vital role in prokaryotic cytokinesis is structurally and functionally very similar to tubulin, the building block of the microtubules that make up the mitotic spindle fibers that are necessary for eukaryotes. FtsZ proteins can form filaments, rings, and other three-dimensional structures that resemble the way tubulin forms microtubules, centrioles, and various cytoskeletal components. In addition, both FtsZ and tubulin employ the same energy source, GTP (guanosine triphosphate), to rapidly assemble and disassemble complex structures. FtsZ and tubulin are homologous structures derived from common evolutionary origins. In this example, FtsZ is the ancestor protein to tubulin (a modern protein). While both proteins are found in extant organisms, tubulin function has evolved and diversified tremendously since evolving from its FtsZ prokaryotic origin. A survey of mitotic assembly components found in present-day unicellular eukaryotes reveals crucial intermediary steps to the complex membrane-enclosed genomes of multicellular eukaryotes ([\[link\]](#)).

**Cell Division Apparatus among Various Organisms**

	<b>Structure of genetic material</b>	<b>Division of nuclear material</b>	<b>Separation of daughter cells</b>

Cell Division Apparatus among Various Organisms			
	Structure of genetic material	Division of nuclear material	Separation of daughter cells
Prokaryotes	There is no nucleus. The single, circular chromosome exists in a region of cytoplasm called the nucleoid.	Occurs through binary fission. As the chromosome is replicated, the two copies move to opposite ends of the cell by an unknown mechanism.	FtsZ proteins assemble into a ring that pinches the cell in two.
Some protists	Linear chromosomes exist in the nucleus.	Chromosomes attach to the nuclear envelope, which remains intact. The mitotic spindle passes through the envelope and elongates the cell. No centrioles exist.	Microfilaments form a cleavage furrow that pinches the cell in two.

Cell Division Apparatus among Various Organisms			
	Structure of genetic material	Division of nuclear material	Separation of daughter cells
Other protists	Linear chromosomes exist in the nucleus.	A mitotic spindle forms from the centrioles and passes through the nuclear membrane, which remains intact. Chromosomes attach to the mitotic spindle, which separates the chromosomes and elongates the cell.	Microfilaments form a cleavage furrow that pinches the cell in two.

Cell Division Apparatus among Various Organisms			
	Structure of genetic material	Division of nuclear material	Separation of daughter cells
Animal cells	Linear chromosomes exist in the nucleus.	A mitotic spindle forms from the centrosomes. The nuclear envelope dissolves. Chromosomes attach to the mitotic spindle, which separates the chromosomes and elongates the cell.	Microfilaments form a cleavage furrow that pinches the cell in two.

## Section Summary

In both prokaryotic and eukaryotic cell division, the genomic DNA is replicated and then each copy is allocated into a daughter cell. In addition, the cytoplasmic contents are divided evenly and distributed to the new cells. However, there are many differences between prokaryotic and eukaryotic cell division. Bacteria have a single, circular DNA chromosome but no nucleus. Therefore, mitosis is not necessary in bacterial cell division. Bacterial cytokinesis is directed by a ring composed of a protein called FtsZ. Ingrowth of membrane and cell wall material from the periphery of the cells

results in the formation of a septum that eventually constructs the separate cell walls of the daughter cells.

## Review Questions

### Exercise:

#### Problem:

Which eukaryotic cell cycle event is missing in binary fission?

- a. cell growth
- b. DNA duplication
- c. karyokinesis
- d. cytokinesis

---

#### Solution:

C

### Exercise:

#### Problem:

FtsZ proteins direct the formation of a \_\_\_\_\_ that will eventually form the new cell walls of the daughter cells.

- a. contractile ring
- b. cell plate
- c. cytoskeleton
- d. septum

---

#### Solution:

B

## Free Response

## **Exercise:**

### **Problem:**

Name the common components of eukaryotic cell division and binary fission.

---

### **Solution:**

The common components of eukaryotic cell division and binary fission are DNA duplication, segregation of duplicated chromosomes, and division of the cytoplasmic contents.

## **Exercise:**

### **Problem:**

Describe how the duplicated bacterial chromosomes are distributed into new daughter cells without the direction of the mitotic spindle.

---

### **Solution:**

As the chromosome is being duplicated, each origin moves away from the starting point of replication. The chromosomes are attached to the cell membrane via proteins; the growth of the membrane as the cell elongates aids in their movement.

## **Glossary**

binary fission

prokaryotic cell division process

FtsZ

tubulin-like protein component of the prokaryotic cytoskeleton that is important in prokaryotic cytokinesis (name origin: **F**ilamenting **t**emperature-sensitive mutant **Z**)

origin

(also, ORI) region of the prokaryotic chromosome where replication begins (origin of replication)

septum

structure formed in a bacterial cell as a precursor to the separation of the cell into two daughter cells



## Introduction

class="introduction"

Each of us,  
like these  
other large  
multicellula  
r organisms,  
begins life  
as a  
fertilized  
egg. After  
trillions of  
cell  
divisions,  
each of us  
develops  
into a  
complex,  
multicellula  
r organism.  
(credit a:  
modificatio  
n of work  
by Frank  
Wouters;  
credit b:  
modificatio  
n of work  
by Ken  
Cole,  
USGS;  
credit c:  
modificatio  
n of work  
by Martin  
Pettitt)



(a)

(b)

(c)

The ability to reproduce *in kind* is a basic characteristic of all living things. *In kind* means that the offspring of any organism closely resemble their parent or parents. Hippopotamuses give birth to hippopotamus calves, Joshua trees produce seeds from which Joshua tree seedlings emerge, and adult flamingos lay eggs that hatch into flamingo chicks. *In kind* does not generally mean *exactly the same*. Whereas many unicellular organisms and a few multicellular organisms can produce genetically identical clones of themselves through cell division, many single-celled organisms and most multicellular organisms reproduce regularly using another method. Sexual reproduction is the production by parents of two haploid cells and the fusion of two haploid cells to form a single, unique diploid cell. In most plants and animals, through tens of rounds of mitotic cell division, this diploid cell will develop into an adult organism. Haploid cells that are part of the sexual reproductive cycle are produced by a type of cell division called meiosis. Sexual reproduction, specifically meiosis and fertilization, introduces variation into offspring that may account for the evolutionary success of sexual reproduction. The vast majority of eukaryotic organisms, both multicellular and unicellular, can or must employ some form of meiosis and fertilization to reproduce.

## The Process of Meiosis

By the end of this section, you will be able to:

- Describe the behavior of chromosomes during meiosis
- Describe cellular events during meiosis
- Explain the differences between meiosis and mitosis
- Explain the mechanisms within meiosis that generate genetic variation among the products of meiosis

Sexual reproduction requires **fertilization**, the union of two cells from two individual organisms. If those two cells each contain one set of chromosomes, then the resulting cell contains two sets of chromosomes. Haploid cells contain one set of chromosomes. Cells containing two sets of chromosomes are called diploid. The number of sets of chromosomes in a cell is called its ploidy level. If the reproductive cycle is to continue, then the diploid cell must somehow reduce its number of chromosome sets before fertilization can occur again, or there will be a continual doubling in the number of chromosome sets in every generation. So, in addition to fertilization, sexual reproduction includes a nuclear division that reduces the number of chromosome sets.

Most animals and plants are diploid, containing two sets of chromosomes. In each **somatic cell** of the organism (all cells of a multicellular organism except the gametes or reproductive cells), the nucleus contains two copies of each chromosome, called homologous chromosomes. Somatic cells are sometimes referred to as “body” cells. Homologous chromosomes are matched pairs containing the same genes in identical locations along their length. Diploid organisms inherit one copy of each homologous chromosome from each parent; all together, they are considered a full set of chromosomes. Haploid cells, containing a single copy of each homologous chromosome, are found only within structures that give rise to either gametes or spores. **Spores** are haploid cells that can produce a haploid organism or can fuse with another spore to form a diploid cell. All animals and most plants produce eggs and sperm, or gametes. Some plants and all fungi produce spores.

The nuclear division that forms haploid cells, which is called **meiosis**, is related to mitosis. As you have learned, mitosis is the part of a cell

reproduction cycle that results in identical daughter nuclei that are also genetically identical to the original parent nucleus. In mitosis, both the parent and the daughter nuclei are at the same ploidy level—diploid for most plants and animals. Meiosis employs many of the same mechanisms as mitosis. However, the starting nucleus is always diploid and the nuclei that result at the end of a meiotic cell division are haploid. To achieve this reduction in chromosome number, meiosis consists of one round of chromosome duplication and two rounds of nuclear division. Because the events that occur during each of the division stages are analogous to the events of mitosis, the same stage names are assigned. However, because there are two rounds of division, the major process and the stages are designated with a “I” or a “II.” Thus, **meiosis I** is the first round of meiotic division and consists of prophase I, prometaphase I, and so on. **Meiosis II**, in which the second round of meiotic division takes place, includes prophase II, prometaphase II, and so on.

## **Meiosis I**

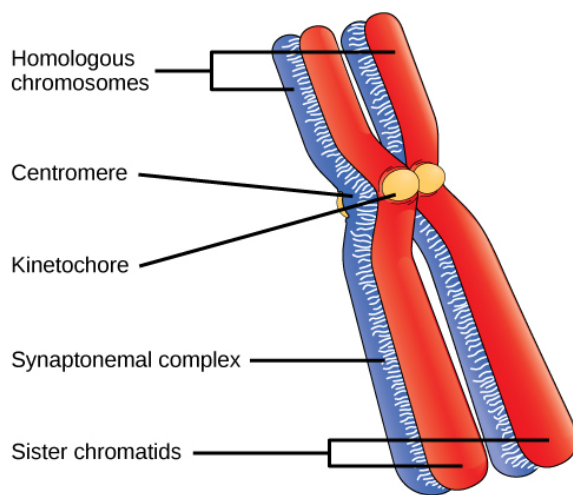
Meiosis is preceded by an interphase consisting of the  $G_1$ , S, and  $G_2$  phases, which are nearly identical to the phases preceding mitosis. The  $G_1$  phase, which is also called the first gap phase, is the first phase of the interphase and is focused on cell growth. The S phase is the second phase of interphase, during which the DNA of the chromosomes is replicated. Finally, the  $G_2$  phase, also called the second gap phase, is the third and final phase of interphase; in this phase, the cell undergoes the final preparations for meiosis.

During DNA duplication in the S phase, each chromosome is replicated to produce two identical copies, called sister chromatids, that are held together at the centromere by **cohesin** proteins. Cohesin holds the chromatids together until anaphase II. The centrosomes, which are the structures that organize the microtubules of the meiotic spindle, also replicate. This prepares the cell to enter prophase I, the first meiotic phase.

## **Prophase I**

Early in prophase I, before the chromosomes can be seen clearly microscopically, the homologous chromosomes are attached at their tips to the nuclear envelope by proteins. As the nuclear envelope begins to break down, the proteins associated with homologous chromosomes bring the pair close to each other. Recall that, in mitosis, homologous chromosomes do not pair together. In mitosis, homologous chromosomes line up end-to-end so that when they divide, each daughter cell receives a sister chromatid from both members of the homologous pair. The **synaptonemal complex**, a lattice of proteins between the homologous chromosomes, first forms at specific locations and then spreads to cover the entire length of the chromosomes. The tight pairing of the homologous chromosomes is called **synapsis**. In synapsis, the genes on the chromatids of the homologous chromosomes are aligned precisely with each other. The synaptonemal complex supports the exchange of chromosomal segments between non-sister homologous chromatids, a process called crossing over. Crossing over can be observed visually after the exchange as **chiasmata** (singular = chiasma) ([link](#)).

In species such as humans, even though the X and Y sex chromosomes are not homologous (most of their genes differ), they have a small region of homology that allows the X and Y chromosomes to pair up during prophase I. A partial synaptonemal complex develops only between the regions of homology.

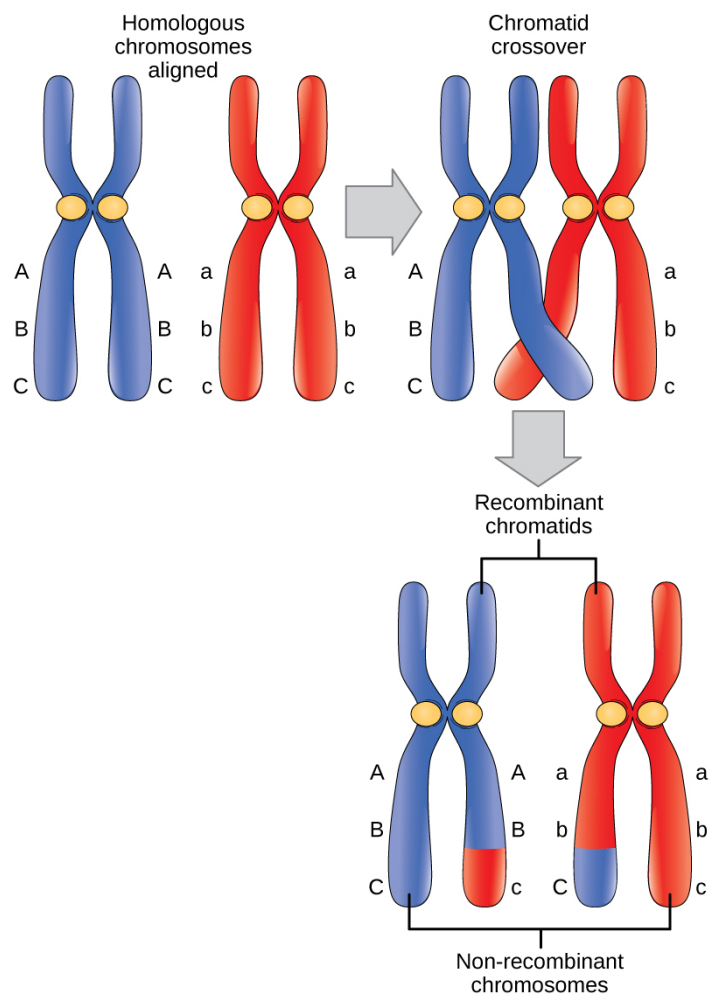


Early in prophase I, homologous chromosomes come together to form a synapse. The chromosomes are bound tightly together and in perfect alignment by a protein lattice called a synaptonemal complex and by cohesin proteins at the centromere.

Located at intervals along the synaptonemal complex are large protein assemblies called **recombination nodules**. These assemblies mark the points of later chiasmata and mediate the multistep process of **crossover**—or genetic recombination—between the non-sister chromatids. Near the recombination nodule on each chromatid, the double-stranded DNA is cleaved, the cut ends are modified, and a new connection is made between the non-sister chromatids. As prophase I progresses, the synaptonemal complex begins to break down and the chromosomes begin to condense. When the synaptonemal complex is gone, the homologous chromosomes remain attached to each other at the centromere and at chiasmata. The chiasmata remain until anaphase I. The number of chiasmata varies according to the species and the length of the chromosome. There must be at least one chiasma per chromosome for proper separation of homologous chromosomes during meiosis I, but there may be as many as 25. Following crossover, the synaptonemal complex breaks down and the cohesin connection between homologous pairs is also removed. At the end of prophase I, the pairs are held together only at the chiasmata ([\[link\]](#)) and are called **tetrads** because the four sister chromatids of each pair of homologous chromosomes are now visible.

The crossover events are the first source of genetic variation in the nuclei produced by meiosis. A single crossover event between homologous non-sister chromatids leads to a reciprocal exchange of equivalent DNA between a maternal chromosome and a paternal chromosome. Now, when that sister chromatid is moved into a gamete cell it will carry some DNA from one parent of the individual and some DNA from the other parent. The

sister recombinant chromatid has a combination of maternal and paternal genes that did not exist before the crossover. Multiple crossovers in an arm of the chromosome have the same effect, exchanging segments of DNA to create recombinant chromosomes.



Crossover occurs between non-sister chromatids of homologous chromosomes. The result is an exchange of genetic material between homologous chromosomes.

## **Prometaphase I**

The key event in prometaphase I is the attachment of the spindle fiber microtubules to the kinetochore proteins at the centromeres. Kinetochore proteins are multiprotein complexes that bind the centromeres of a chromosome to the microtubules of the mitotic spindle. Microtubules grow from centrosomes placed at opposite poles of the cell. The microtubules move toward the middle of the cell and attach to one of the two fused homologous chromosomes. The microtubules attach at each chromosome's kinetochores. With each member of the homologous pair attached to opposite poles of the cell, in the next phase, the microtubules can pull the homologous pair apart. A spindle fiber that has attached to a kinetochore is called a kinetochore microtubule. At the end of prometaphase I, each tetrad is attached to microtubules from both poles, with one homologous chromosome facing each pole. The homologous chromosomes are still held together at chiasmata. In addition, the nuclear membrane has broken down entirely.

## **Metaphase I**

During metaphase I, the homologous chromosomes are arranged in the center of the cell with the kinetochores facing opposite poles. The homologous pairs orient themselves randomly at the equator. For example, if the two homologous members of chromosome 1 are labeled a and b, then the chromosomes could line up a-b, or b-a. This is important in determining the genes carried by a gamete, as each will only receive one of the two homologous chromosomes. Recall that homologous chromosomes are not identical. They contain slight differences in their genetic information, causing each gamete to have a unique genetic makeup.

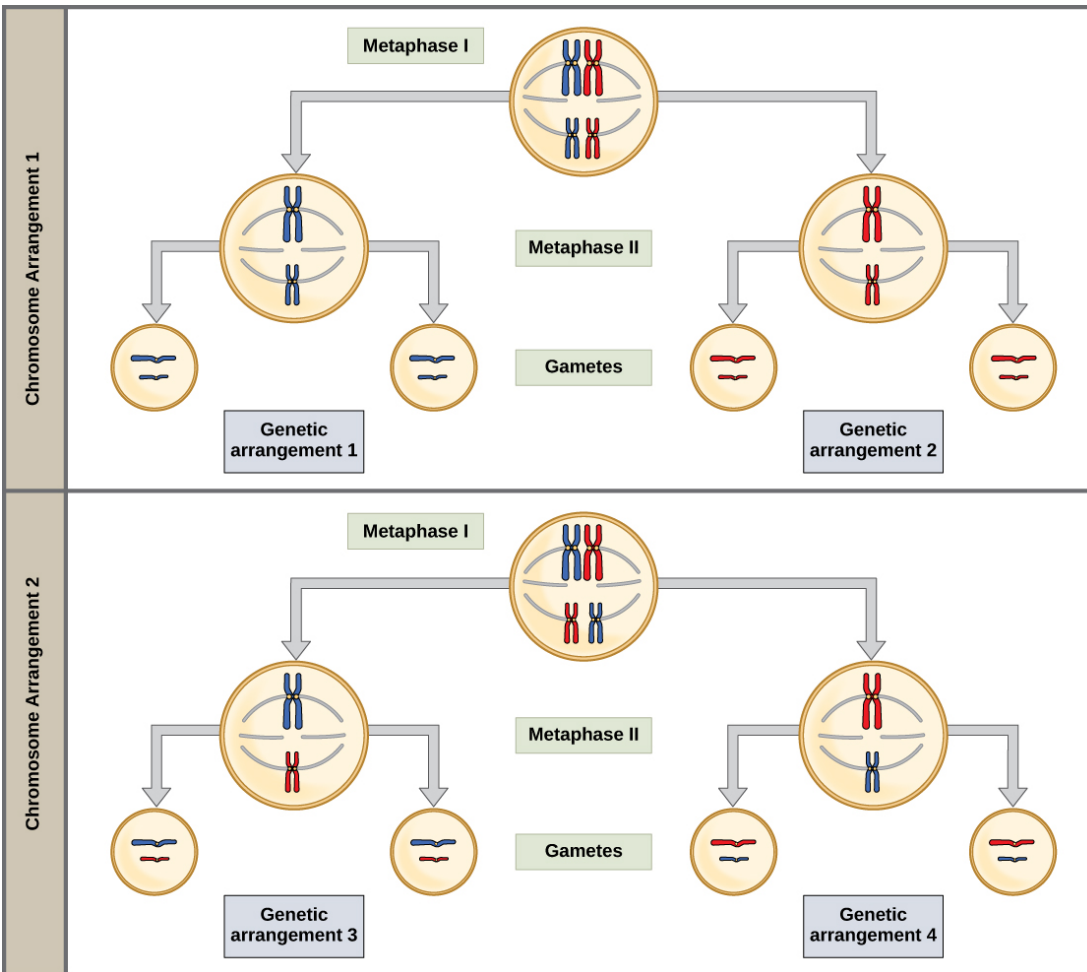
This randomness is the physical basis for the creation of the second form of genetic variation in offspring. Consider that the homologous chromosomes of a sexually reproducing organism are originally inherited as two separate sets, one from each parent. Using humans as an example, one set of 23 chromosomes is present in the egg donated by the mother. The father provides the other set of 23 chromosomes in the sperm that fertilizes the



egg. Every cell of the multicellular offspring has copies of the original two sets of homologous chromosomes. In prophase I of meiosis, the homologous chromosomes form the tetrads. In metaphase I, these pairs line up at the midway point between the two poles of the cell to form the metaphase plate. Because there is an equal chance that a microtubule fiber will encounter a maternally or paternally inherited chromosome, the arrangement of the tetrads at the metaphase plate is random. Any maternally inherited chromosome may face either pole. Any paternally inherited chromosome may also face either pole. The orientation of each tetrad is independent of the orientation of the other 22 tetrads.

This event—the random (or independent) assortment of homologous chromosomes at the metaphase plate—is the second mechanism that introduces variation into the gametes or spores. In each cell that undergoes meiosis, the arrangement of the tetrads is different. The number of variations is dependent on the number of chromosomes making up a set. There are two possibilities for orientation at the metaphase plate; the possible number of alignments therefore equals  $2n$ , where  $n$  is the number of chromosomes per set. Humans have 23 chromosome pairs, which results in over eight million ( $2^{23}$ ) possible genetically-distinct gametes. This number does not include the variability that was previously created in the sister chromatids by crossover. Given these two mechanisms, it is highly unlikely that any two haploid cells resulting from meiosis will have the same genetic composition ([link](#)).

To summarize the genetic consequences of meiosis I, the maternal and paternal genes are recombined by crossover events that occur between each homologous pair during prophase I. In addition, the random assortment of tetrads on the metaphase plate produces a unique combination of maternal and paternal chromosomes that will make their way into the gametes.



Random, independent assortment during metaphase I can be demonstrated by considering a cell with a set of two chromosomes ( $n = 2$ ). In this case, there are two possible arrangements at the equatorial plane in metaphase I. The total possible number of different gametes is  $2n$ , where  $n$  equals the number of chromosomes in a set. In this example, there are four possible genetic combinations for the gametes. With  $n = 23$  in human cells, there are over 8 million possible combinations of paternal and maternal chromosomes.

## Anaphase I

In anaphase I, the microtubules pull the linked chromosomes apart. The sister chromatids remain tightly bound together at the centromere. The chiasmata are broken in anaphase I as the microtubules attached to the fused kinetochores pull the homologous chromosomes apart ([link](#)).

## **Telophase I and Cytokinesis**

In telophase, the separated chromosomes arrive at opposite poles. The remainder of the typical telophase events may or may not occur, depending on the species. In some organisms, the chromosomes decondense and nuclear envelopes form around the chromatids in telophase I. In other organisms, cytokinesis—the physical separation of the cytoplasmic components into two daughter cells—occurs without reformation of the nuclei. In nearly all species of animals and some fungi, cytokinesis separates the cell contents via a cleavage furrow (constriction of the actin ring that leads to cytoplasmic division). In plants, a cell plate is formed during cell cytokinesis by Golgi vesicles fusing at the metaphase plate. This cell plate will ultimately lead to the formation of cell walls that separate the two daughter cells.

Two haploid cells are the end result of the first meiotic division. The cells are haploid because at each pole, there is just one of each pair of the homologous chromosomes. Therefore, only one full set of the chromosomes is present. This is why the cells are considered haploid—there is only one chromosome set, even though each homolog still consists of two sister chromatids. Recall that sister chromatids are merely duplicates of one of the two homologous chromosomes (except for changes that occurred during crossing over). In meiosis II, these two sister chromatids will separate, creating four haploid daughter cells.

### **Note:**

[Link to Learning](#)



Review the process of meiosis, observing how chromosomes align and migrate, at [Meiosis: An Interactive Animation](#).

## Meiosis II

In some species, cells enter a brief interphase, or **interkinesis**, before entering meiosis II. Interkinesis lacks an S phase, so chromosomes are not duplicated. The two cells produced in meiosis I go through the events of meiosis II in synchrony. During meiosis II, the sister chromatids within the two daughter cells separate, forming four new haploid gametes. The mechanics of meiosis II is similar to mitosis, except that each dividing cell has only one set of homologous chromosomes. Therefore, each cell has half the number of sister chromatids to separate out as a diploid cell undergoing mitosis.

### Prophase II

If the chromosomes decondensed in telophase I, they condense again. If nuclear envelopes were formed, they fragment into vesicles. The centrosomes that were duplicated during interkinesis move away from each other toward opposite poles, and new spindles are formed.

### Prometaphase II

The nuclear envelopes are completely broken down, and the spindle is fully formed. Each sister chromatid forms an individual kinetochore that attaches

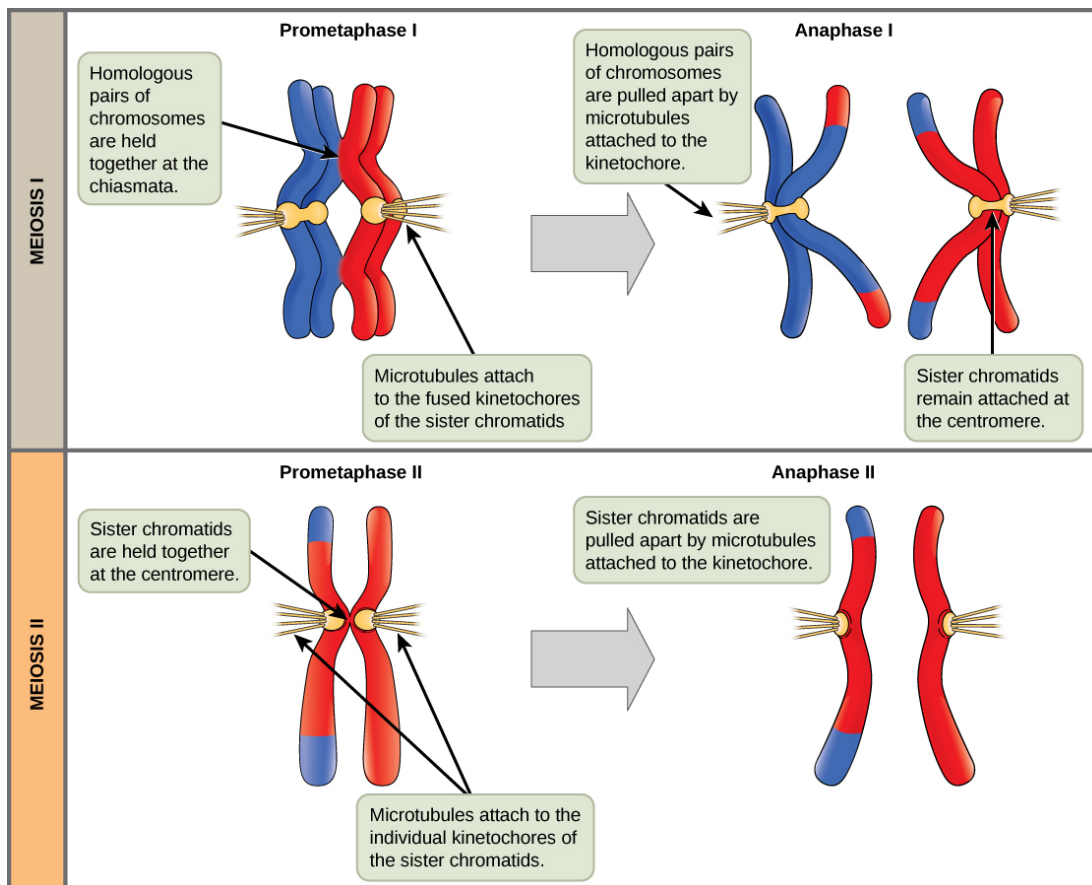
to microtubules from opposite poles.

## Metaphase II

The sister chromatids are maximally condensed and aligned at the equator of the cell.

## Anaphase II

The sister chromatids are pulled apart by the kinetochore microtubules and move toward opposite poles. Non-kinetochore microtubules elongate the cell.



The process of chromosome alignment differs between meiosis I and meiosis II. In prometaphase I, microtubules attach to the fused kinetochores of homologous chromosomes, and the homologous chromosomes are arranged at the midpoint of the cell in metaphase I. In anaphase I, the homologous chromosomes are separated. In prometaphase II, microtubules attach to the kinetochores of sister chromatids, and the sister chromatids are arranged at the midpoint of the cells in metaphase II. In anaphase II, the sister chromatids are separated.

### **Telophase II and Cytokinesis**

The chromosomes arrive at opposite poles and begin to decondense. Nuclear envelopes form around the chromosomes. Cytokinesis separates the two cells into four unique haploid cells. At this point, the newly formed nuclei are both haploid. The cells produced are genetically unique because of the random assortment of paternal and maternal homologs and because of the recombining of maternal and paternal segments of chromosomes (with their sets of genes) that occurs during crossover. The entire process of meiosis is outlined in [\[link\]](#).

Stage	Event	Outcome
INTERPHASE	<b>S phase</b> Nuclear envelope Centrosomes (with centriole pairs) Chromatin	Chromosomes are duplicated during interphase. The resulting sister chromatids are held together at the centromere. The centrosomes are also duplicated.
	<b>Prophase I</b> Sister chromatids Spindle Chiasmata Tetrad	Chromosomes condense, and the nuclear envelope fragments. Homologous chromosomes bind firmly together along their length, forming a tetrad. Chiasmata form between non-sister chromatids. Crossing over occurs at the chiasmata. Spindle fibers emerge from the centrosomes.
MEIOSIS I	<b>Prometaphase I</b> Centromere (with kinetochore)	Homologous chromosomes are attached to spindle microtubules at the fused kinetochore shared by the sister chromatids. Chromosomes continue to condense, and the nuclear envelope completely disappears.
	<b>Metaphase I</b> Microtubule attached to kinetochore Metaphase plate	Homologous chromosomes randomly assemble at the metaphase plate, where they have been maneuvered into place by the microtubules.
	<b>Anaphase I</b> Sister chromatids remain attached Homologous chromosomes separate	Spindle microtubules pull the homologous chromosomes apart. The sister chromatids are still attached at the centromere.
	<b>Telophase I and Cytokinesis</b> Cleavage furrow	Sister chromatids arrive at the poles of the cell and begin to decondense. A nuclear envelope forms around each nucleus and the cytoplasm is divided by a cleavage furrow. The result is two haploid cells. Each cell contains one duplicated copy of each homologous chromosome pair.
MEIOSIS II	<b>Prophase II</b>	Sister chromatids condense. A new spindle begins to form. The nuclear envelope starts to fragment.
	<b>Prometaphase II</b>	The nuclear envelope disappears, and the spindle fibers engage the individual kinetochores on the sister chromatids.
	<b>Metaphase II</b>	Sister chromatids line up at the metaphase plate.
	<b>Anaphase II</b> Sister chromatids separate	Sister chromatids are pulled apart by the shortening of the kinetochore microtubules. Non-kinetochore microtubules lengthen the cell.
	<b>Telophase II and Cytokinesis</b> Haploid daughter cells	Chromosomes arrive at the poles of the cell and decondense. Nuclear envelopes surround the four nuclei. Cleavage furrows divide the two cells into four haploid cells.

An animal cell with a diploid number of four ( $2n = 4$ ) proceeds through the stages of meiosis to form four haploid daughter cells.

## Comparing Meiosis and Mitosis

Mitosis and meiosis are both forms of division of the nucleus in eukaryotic cells. They share some similarities, but also exhibit distinct differences that lead to very different outcomes ([link](#)). Mitosis is a single nuclear division that results in two nuclei that are usually partitioned into two new cells. The nuclei resulting from a mitotic division are genetically identical to the original nucleus. They have the same number of sets of chromosomes, one set in the case of haploid cells and two sets in the case of diploid cells. In most plants and all animal species, it is typically diploid cells that undergo mitosis to form new diploid cells. In contrast, meiosis consists of two nuclear divisions resulting in four nuclei that are usually partitioned into four new cells. The nuclei resulting from meiosis are not genetically identical and they contain one chromosome set only. This is half the number of chromosome sets in the original cell, which is diploid.

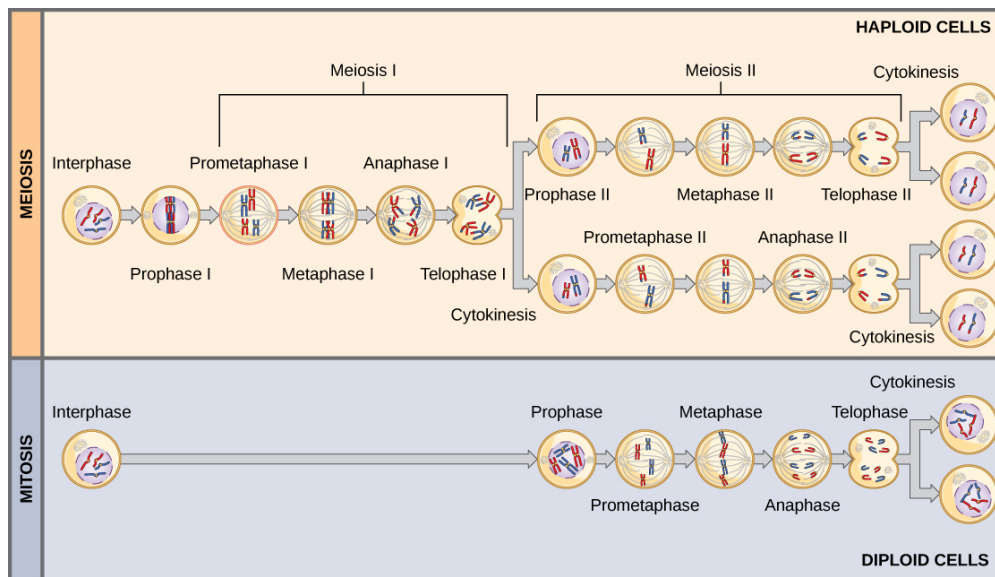
The main differences between mitosis and meiosis occur in meiosis I, which is a very different nuclear division than mitosis. In meiosis I, the homologous chromosome pairs become associated with each other, are bound together with the synaptonemal complex, develop chiasmata and undergo crossover between sister chromatids, and line up along the metaphase plate in tetrads with kinetochore fibers from opposite spindle poles attached to each kinetochore of a homolog in a tetrad. All of these events occur only in meiosis I.

When the chiasmata resolve and the tetrad is broken up with the homologs moving to one pole or another, the ploidy level—the number of sets of chromosomes in each future nucleus—has been reduced from two to one. For this reason, meiosis I is referred to as a **reduction division**. There is no such reduction in ploidy level during mitosis.

Meiosis II is much more analogous to a mitotic division. In this case, the duplicated chromosomes (only one set of them) line up on the metaphase plate with divided kinetochores attached to kinetochore fibers from opposite poles. During anaphase II, as in mitotic anaphase, the kinetochores divide and one sister chromatid—now referred to as a chromosome—is pulled to one pole while the other sister chromatid is pulled to the other pole. If it



were not for the fact that there had been crossover, the two products of each individual meiosis II division would be identical (like in mitosis). Instead, they are different because there has always been at least one crossover per chromosome. Meiosis II is not a reduction division because although there are fewer copies of the genome in the resulting cells, there is still one set of chromosomes, as there was at the end of meiosis I.



						OUTCOME
PROCESS	DNA synthesis	Synapsis of homologous chromosomes	Crossover	Homologous chromosomes line up at metaphase plate	Sister chromatids line up at metaphase plate	Number and genetic composition of daughter cells
MEIOSIS	Occurs in S phase of interphase	During prophase I	During prophase I	During metaphase I	During metaphase II	Four haploid cells at the end of meiosis II
MITOSIS	Occurs in S phase of interphase	Does not occur in mitosis	Does not occur in mitosis	Does not occur in mitosis	During metaphase	Two diploid cells at the end of mitosis

Meiosis and mitosis are both preceded by one round of DNA replication; however, meiosis includes two nuclear divisions. The four daughter cells resulting from meiosis are haploid and genetically distinct. The daughter cells resulting from mitosis are diploid and identical to the parent cell.

**Note:****Evolution Connection****The Mystery of the Evolution of Meiosis**

Some characteristics of organisms are so widespread and fundamental that it is sometimes difficult to remember that they evolved like other simpler traits. Meiosis is such an extraordinarily complex series of cellular events that biologists have had trouble hypothesizing and testing how it may have evolved. Although meiosis is inextricably entwined with sexual reproduction and its advantages and disadvantages, it is important to separate the questions of the evolution of meiosis and the evolution of sex, because early meiosis may have been advantageous for different reasons than it is now. Thinking outside the box and imagining what the early benefits from meiosis might have been is one approach to uncovering how it may have evolved.

Meiosis and mitosis share obvious cellular processes and it makes sense that meiosis evolved from mitosis. The difficulty lies in the clear differences between meiosis I and mitosis. Adam Wilkins and Robin Holliday<sup>[footnote]</sup> summarized the unique events that needed to occur for the evolution of meiosis from mitosis. These steps are homologous chromosome pairing, crossover exchanges, sister chromatids remaining attached during anaphase, and suppression of DNA replication in interphase. They argue that the first step is the hardest and most important, and that understanding how it evolved would make the evolutionary process clearer. They suggest genetic experiments that might shed light on the evolution of synapsis.

Adam S. Wilkins and Robin Holliday, “The Evolution of Meiosis from Mitosis,” *Genetics* 181 (2009): 3–12.

There are other approaches to understanding the evolution of meiosis in progress. Different forms of meiosis exist in single-celled protists. Some appear to be simpler or more “primitive” forms of meiosis. Comparing the meiotic divisions of different protists may shed light on the evolution of meiosis. Marilee Ramesh and colleagues<sup>[footnote]</sup> compared the genes involved in meiosis in protists to understand when and where meiosis might have evolved. Although research is still ongoing, recent scholarship

into meiosis in protists suggests that some aspects of meiosis may have evolved later than others. This kind of genetic comparison can tell us what aspects of meiosis are the oldest and what cellular processes they may have borrowed from in earlier cells.

Marilee A. Ramesh, Shehre-Banoo Malik and John M. Logsdon, Jr, “A Phylogenetic Inventory of Meiotic Genes: Evidence for Sex in *Giardia* and an Early Eukaryotic Origin of Meiosis,” *Current Biology* 15 (2005):185–91.

**Note:**

Link to Learning



Click through the steps of this interactive animation to compare the meiotic process of cell division to that of mitosis: [How Cells Divide](#).

## Section Summary

Sexual reproduction requires that diploid organisms produce haploid cells that can fuse during fertilization to form diploid offspring. As with mitosis, DNA replication occurs prior to meiosis during the S-phase of the cell cycle. Meiosis is a series of events that arrange and separate chromosomes and chromatids into daughter cells. During the interphases of meiosis, each chromosome is duplicated. In meiosis, there are two rounds of nuclear division resulting in four nuclei and usually four daughter cells, each with half the number of chromosomes as the parent cell. The first separates homologs, and the second—like mitosis—separates chromatids into individual chromosomes. During meiosis, variation in the daughter nuclei is

introduced because of crossover in prophase I and random alignment of tetrads at metaphase I. The cells that are produced by meiosis are genetically unique.

Meiosis and mitosis share similarities, but have distinct outcomes. Mitotic divisions are single nuclear divisions that produce daughter nuclei that are genetically identical and have the same number of chromosome sets as the original cell. Meiotic divisions include two nuclear divisions that produce four daughter nuclei that are genetically different and have one chromosome set instead of the two sets of chromosomes in the parent cell. The main differences between the processes occur in the first division of meiosis, in which homologous chromosomes are paired and exchange non-sister chromatid segments. The homologous chromosomes separate into different nuclei during meiosis I, causing a reduction of ploidy level in the first division. The second division of meiosis is more similar to a mitotic division, except that the daughter cells do not contain identical genomes because of crossover.

## Review Questions

### Exercise:

**Problem:** Meiosis produces \_\_\_\_\_ daughter cells.

- a. two haploid
- b. two diploid
- c. four haploid
- d. four diploid

---

### Solution:

C

### Exercise:

**Problem:** What structure is most important in forming the tetrads?

- a. centromere
- b. synaptonemal complex
- c. chiasma
- d. kinetochore

---

**Solution:**

B

**Exercise:**

**Problem:**

At which stage of meiosis are sister chromatids separated from each other?

- a. prophase I
- b. prophase II
- c. anaphase I
- d. anaphase II

---

**Solution:**

D

**Exercise:**

**Problem:**

At metaphase I, homologous chromosomes are connected only at what structures?

- a. chiasmata
- b. recombination nodules
- c. microtubules
- d. kinetochores

---

**Solution:**

A

**Exercise:**

**Problem:** Which of the following is *not* true in regard to crossover?

- a. Spindle microtubules guide the transfer of DNA across the synaptonemal complex.
- b. Non-sister chromatids exchange genetic material.
- c. Chiasmata are formed.
- d. Recombination nodules mark the crossover point.

---

**Solution:**

C

**Exercise:**

**Problem:**

What phase of mitotic interphase is missing from meiotic interkinesis?

- a. G<sub>0</sub> phase
- b. G<sub>1</sub> phase
- c. S phase
- d. G<sub>2</sub> phase

---

**Solution:**

C

**Exercise:**

**Problem:** The part of meiosis that is similar to mitosis is \_\_\_\_\_.

- a. meiosis I
- b. anaphase I
- c. meiosis II

d. interkinesis

---

**Solution:**

C

**Exercise:**

**Problem:**

If a muscle cell of a typical organism has 32 chromosomes, how many chromosomes will be in a gamete of that same organism?

- a. 8
- b. 16
- c. 32
- d. 64

---

**Solution:**

B

**Free Response**

**Exercise:**

**Problem:** Describe the process that results in the formation of a tetrad.

---

**Solution:**

During the meiotic interphase, each chromosome is duplicated. The sister chromatids that are formed during synthesis are held together at the centromere region by cohesin proteins. All chromosomes are attached to the nuclear envelope by their tips. As the cell enters prophase I, the nuclear envelope begins to fragment, and the proteins holding homologous chromosomes locate each other. The four sister

chromatids align lengthwise, and a protein lattice called the synaptonemal complex is formed between them to bind them together. The synaptonemal complex facilitates crossover between non-sister chromatids, which is observed as chiasmata along the length of the chromosome. As prophase I progresses, the synaptonemal complex breaks down and the sister chromatids become free, except where they are attached by chiasmata. At this stage, the four chromatids are visible in each homologous pairing and are called a tetrad.

**Exercise:**

**Problem:**

Explain how the random alignment of homologous chromosomes during metaphase I contributes to the variation in gametes produced by meiosis.

---

**Solution:**

Random alignment leads to new combinations of traits. The chromosomes that were originally inherited by the gamete-producing individual came equally from the egg and the sperm. In metaphase I, the duplicated copies of these maternal and paternal homologous chromosomes line up across the center of the cell. The orientation of each tetrad is random. There is an equal chance that the maternally derived chromosomes will be facing either pole. The same is true of the paternally derived chromosomes. The alignment should occur differently in almost every meiosis. As the homologous chromosomes are pulled apart in anaphase I, any combination of maternal and paternal chromosomes will move toward each pole. The gametes formed from these two groups of chromosomes will have a mixture of traits from the individual's parents. Each gamete is unique.

**Exercise:**

**Problem:**

What is the function of the fused kinetochore found on sister chromatids in prometaphase I?



---

**Solution:**

In metaphase I, the homologous chromosomes line up at the metaphase plate. In anaphase I, the homologous chromosomes are pulled apart and move to opposite poles. Sister chromatids are not separated until meiosis II. The fused kinetochore formed during meiosis I ensures that each spindle microtubule that binds to the tetrad will attach to both sister chromatids.

**Exercise:****Problem:**

In a comparison of the stages of meiosis to the stages of mitosis, which stages are unique to meiosis and which stages have the same events in both meiosis and mitosis?

---

**Solution:**

All of the stages of meiosis I, except possibly telophase I, are unique because homologous chromosomes are separated, not sister chromatids. In some species, the chromosomes do not decondense and the nuclear envelopes do not form in telophase I. All of the stages of meiosis II have the same events as the stages of mitosis, with the possible exception of prophase II. In some species, the chromosomes are still condensed and there is no nuclear envelope. Other than this, all processes are the same.

**Glossary****chiasmata**

(singular, *chiasma*) the structure that forms at the crossover points after genetic material is exchanged

**cohesin**

proteins that form a complex that seals sister chromatids together at their centromeres until anaphase II of meiosis

crossover

exchange of genetic material between non-sister chromatids resulting in chromosomes that incorporate genes from both parents of the organism

fertilization

union of two haploid cells from two individual organisms

interkinesis

(also, *interphase II*) brief period of rest between meiosis I and meiosis II

meiosis

a nuclear division process that results in four haploid cells

meiosis I

first round of meiotic cell division; referred to as reduction division because the ploidy level is reduced from diploid to haploid

meiosis II

second round of meiotic cell division following meiosis I; sister chromatids are separated into individual chromosomes, and the result is four unique haploid cells

recombination nodules

protein assemblies formed on the synaptonemal complex that mark the points of crossover events and mediate the multistep process of genetic recombination between non-sister chromatids

reduction division

nuclear division that produces daughter nuclei each having one-half as many chromosome sets as the parental nucleus; meiosis I is a reduction division

somatic cell

all the cells of a multicellular organism except the gametes or reproductive cells

spore

haploid cell that can produce a haploid multicellular organism or can fuse with another spore to form a diploid cell

synapsis

formation of a close association between homologous chromosomes during prophase I

synaptonemal complex

protein lattice that forms between homologous chromosomes during prophase I, supporting crossover

tetrad

two duplicated homologous chromosomes (four chromatids) bound together by chiasmata during prophase I

## Sexual Reproduction

By the end of this section, you will be able to:

- Explain that meiosis and sexual reproduction are evolved traits
- Identify variation among offspring as a potential evolutionary advantage to sexual reproduction
- Describe the three different life-cycle types among sexual multicellular organisms and their commonalities

Sexual reproduction was an early evolutionary innovation after the appearance of eukaryotic cells. It appears to have been very successful because most eukaryotes are able to reproduce sexually, and in many animals, it is the only mode of reproduction. And yet, scientists recognize some real disadvantages to sexual reproduction. On the surface, creating offspring that are genetic clones of the parent appears to be a better system. If the parent organism is successfully occupying a habitat, offspring with the same traits would be similarly successful. There is also the obvious benefit to an organism that can produce offspring whenever circumstances are favorable by asexual budding, fragmentation, or asexual eggs. These methods of reproduction do not require another organism of the opposite sex. Indeed, some organisms that lead a solitary lifestyle have retained the ability to reproduce asexually. In addition, in asexual populations, every individual is capable of reproduction. In sexual populations, the males are not producing the offspring themselves, so in theory an asexual population could grow twice as fast.

However, multicellular organisms that exclusively depend on asexual reproduction are exceedingly rare. Why is sexuality (and meiosis) so common? This is one of the important unanswered questions in biology and has been the focus of much research beginning in the latter half of the twentieth century. There are several possible explanations, one of which is that the variation that sexual reproduction creates among offspring is very important to the survival and reproduction of the population. Thus, on average, a sexually reproducing population will leave more descendants than an otherwise similar asexually reproducing population. The only source of variation in asexual organisms is mutation. This is the ultimate source of variation in sexual organisms, but in addition, those different

mutations are continually reshuffled from one generation to the next when different parents combine their unique genomes and the genes are mixed into different combinations by crossovers during prophase I and random assortment at metaphase I.

**Note:**

**Evolution Connection**

**The Red Queen Hypothesis**

It is not in dispute that sexual reproduction provides evolutionary advantages to organisms that employ this mechanism to produce offspring. But why, even in the face of fairly stable conditions, does sexual reproduction persist when it is more difficult and costly for individual organisms? Variation is the outcome of sexual reproduction, but why are ongoing variations necessary? Enter the Red Queen hypothesis, first proposed by Leigh Van Valen in 1973. [\[footnote\]](#) The concept was named in reference to the Red Queen's race in Lewis Carroll's book, *Through the Looking-Glass*.

Leigh Van Valen, "A New Evolutionary Law," *Evolutionary Theory* 1 (1973): 1–30

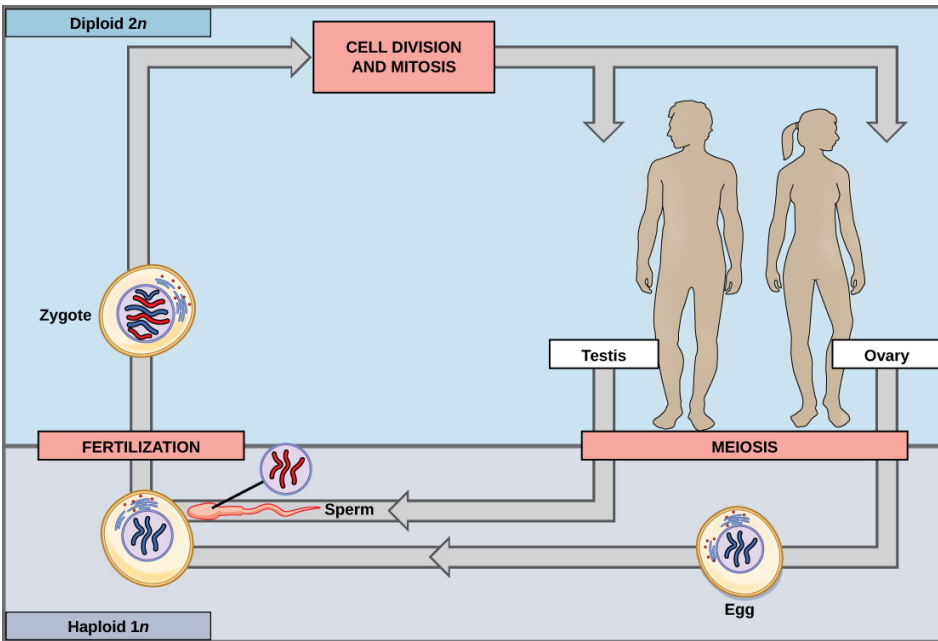
All species co-evolve with other organisms; for example predators evolve with their prey, and parasites evolve with their hosts. Each tiny advantage gained by favorable variation gives a species an edge over close competitors, predators, parasites, or even prey. The only method that will allow a co-evolving species to maintain its own share of the resources is to also continually improve its fitness. As one species gains an advantage, this increases selection on the other species; they must also develop an advantage or they will be outcompeted. No single species progresses too far ahead because genetic variation among the progeny of sexual reproduction provides all species with a mechanism to improve rapidly. Species that cannot keep up become extinct. The Red Queen's catchphrase was, "It takes all the running you can do to stay in the same place." This is an apt description of co-evolution between competing species.

## Life Cycles of Sexually Reproducing Organisms

Fertilization and meiosis alternate in sexual **life cycles**. What happens between these two events depends on the organism. The process of meiosis reduces the chromosome number by half. Fertilization, the joining of two haploid gametes, restores the diploid condition. There are three main categories of life cycles in multicellular organisms: **diploid-dominant**, in which the multicellular diploid stage is the most obvious life stage, such as with most animals including humans; **haploid-dominant**, in which the multicellular haploid stage is the most obvious life stage, such as with all fungi and some algae; and **alternation of generations**, in which the two stages are apparent to different degrees depending on the group, as with plants and some algae.

### Diploid-Dominant Life Cycle

Nearly all animals employ a diploid-dominant life-cycle strategy in which the only haploid cells produced by the organism are the gametes. Early in the development of the embryo, specialized diploid cells, called **germ cells**, are produced within the gonads, such as the testes and ovaries. Germ cells are capable of mitosis to perpetuate the cell line and meiosis to produce gametes. Once the haploid gametes are formed, they lose the ability to divide again. There is no multicellular haploid life stage. Fertilization occurs with the fusion of two gametes, usually from different individuals, restoring the diploid state ([link](#)).



In animals, sexually reproducing adults form haploid gametes from diploid germ cells. Fusion of the gametes gives rise to a fertilized egg cell, or zygote. The zygote will undergo multiple rounds of mitosis to produce a multicellular offspring. The germ cells are generated early in the development of the zygote.

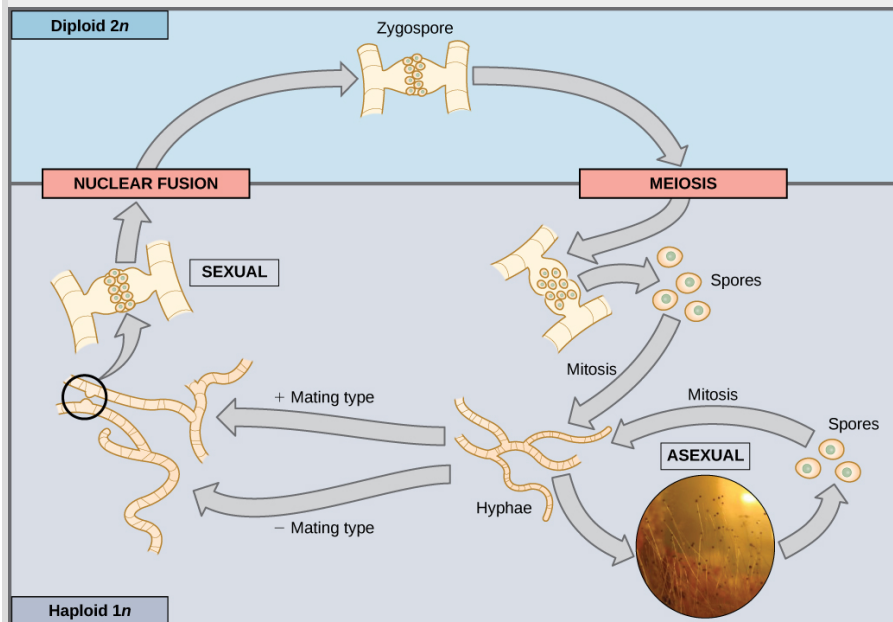
## Haploid-Dominant Life Cycle

Most fungi and algae employ a life-cycle type in which the “body” of the organism—the ecologically important part of the life cycle—is haploid. The haploid cells that make up the tissues of the dominant multicellular stage are formed by mitosis. During sexual reproduction, specialized haploid cells from two individuals, designated the (+) and (−) mating types, join to form a diploid zygote. The zygote immediately undergoes meiosis to form four haploid cells called spores. Although haploid like the “parents,” these spores contain a new genetic combination from two parents. The spores can remain dormant for various time periods. Eventually, when conditions are

conductive, the spores form multicellular haploid structures by many rounds of mitosis ([\[link\]](#)).

**Note:**

**Art Connection**



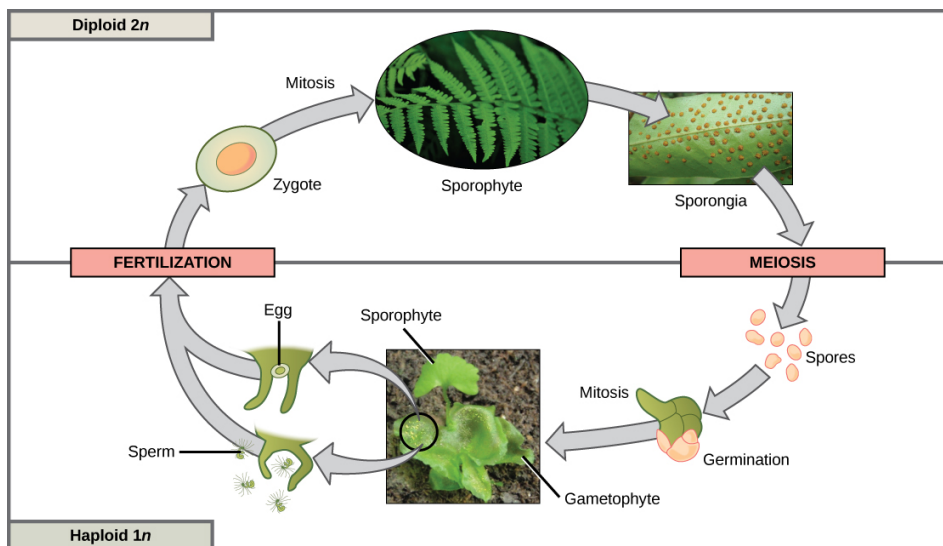
Fungi, such as black bread mold (*Rhizopus nigricans*), have haploid-dominant life cycles. The haploid multicellular stage produces specialized haploid cells by mitosis that fuse to form a diploid zygote. The zygote undergoes meiosis to produce haploid spores. Each spore gives rise to a multicellular haploid organism by mitosis. (credit “zygomycota” micrograph: modification of work by “Fanaberka”/Wikimedia Commons)

If a mutation occurs so that a fungus is no longer able to produce a minus mating type, will it still be able to reproduce?



## Alternation of Generations

The third life-cycle type, employed by some algae and all plants, is a blend of the haploid-dominant and diploid-dominant extremes. Species with alternation of generations have both haploid and diploid multicellular organisms as part of their life cycle. The haploid multicellular plants are called **gametophytes**, because they produce gametes from specialized cells. Meiosis is not directly involved in the production of gametes in this case, because the organism that produces the gametes is already a haploid. Fertilization between the gametes forms a diploid zygote. The zygote will undergo many rounds of mitosis and give rise to a diploid multicellular plant called a **sporophyte**. Specialized cells of the sporophyte will undergo meiosis and produce haploid spores. The spores will subsequently develop into the gametophytes ([\[link\]](#)).



Plants have a life cycle that alternates between a multicellular haploid organism and a multicellular diploid organism. In some plants, such as ferns, both the haploid and diploid plant stages are free-living. The diploid plant is called a sporophyte because it produces haploid spores by meiosis. The spores develop into multicellular, haploid plants called gametophytes because they produce gametes. The

gametes of two individuals will fuse to form a diploid zygote that becomes the sporophyte. (credit “fern”: modification of work by Cory Zanker; credit “sporangia”: modification of work by "Obsidian Soul"/Wikimedia Commons; credit “gametophyte and sporophyte”: modification of work by “Vlmastra”/Wikimedia Commons)

Although all plants utilize some version of the alternation of generations, the relative size of the sporophyte and the gametophyte and the relationship between them vary greatly. In plants such as moss, the gametophyte organism is the free-living plant, and the sporophyte is physically dependent on the gametophyte. In other plants, such as ferns, both the gametophyte and sporophyte plants are free-living; however, the sporophyte is much larger. In seed plants, such as magnolia trees and daisies, the gametophyte is composed of only a few cells and, in the case of the female gametophyte, is completely retained within the sporophyte.

Sexual reproduction takes many forms in multicellular organisms. However, at some point in each type of life cycle, meiosis produces haploid cells that will fuse with the haploid cell of another organism. The mechanisms of variation—crossover, random assortment of homologous chromosomes, and random fertilization—are present in all versions of sexual reproduction. The fact that nearly every multicellular organism on Earth employs sexual reproduction is strong evidence for the benefits of producing offspring with unique gene combinations, though there are other possible benefits as well.

## **Section Summary**

Nearly all eukaryotes undergo sexual reproduction. The variation introduced into the reproductive cells by meiosis appears to be one of the advantages of sexual reproduction that has made it so successful. Meiosis and fertilization alternate in sexual life cycles. The process of meiosis produces unique reproductive cells called gametes, which have half the

number of chromosomes as the parent cell. Fertilization, the fusion of haploid gametes from two individuals, restores the diploid condition. Thus, sexually reproducing organisms alternate between haploid and diploid stages. However, the ways in which reproductive cells are produced and the timing between meiosis and fertilization vary greatly. There are three main categories of life cycles: diploid-dominant, demonstrated by most animals; haploid-dominant, demonstrated by all fungi and some algae; and the alternation of generations, demonstrated by plants and some algae.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) If a mutation occurs so that a fungus is no longer able to produce a minus mating type, will it still be able to reproduce?

---

#### Solution:

[\[link\]](#) Yes, it will be able to reproduce asexually.

## Review Questions

### Exercise:

#### Problem:

What is a likely evolutionary advantage of sexual reproduction over asexual reproduction?

- a. Sexual reproduction involves fewer steps.
- b. There is a lower chance of using up the resources in a given environment.
- c. Sexual reproduction results in variation in the offspring.
- d. Sexual reproduction is more cost-effective.

---

**Solution:**

C

**Exercise:**

**Problem:**

Which type of life cycle has both a haploid and diploid multicellular stage?

- a. asexual
- b. diploid-dominant
- c. haploid-dominant
- d. alternation of generations

---

**Solution:**

D

**Exercise:**

**Problem:** Fungi typically display which type of life cycle?

- a. diploid-dominant
- b. haploid-dominant
- c. alternation of generations
- d. asexual

---

**Solution:**

B

**Exercise:**

**Problem:**

A diploid, multicellular life-cycle stage that gives rise to haploid cells by meiosis is called a \_\_\_\_\_.

- a. sporophyte
- b. gametophyte
- c. spore
- d. gamete

---

**Solution:**

A

**Free Response****Exercise:****Problem:**

List and briefly describe the three processes that lead to variation in offspring with the same parents.

---

**Solution:**

a. Crossover occurs in prophase I between non-sister homologous chromosomes. Segments of DNA are exchanged between maternally derived and paternally derived chromosomes, and new gene combinations are formed. b. Random alignment during metaphase I leads to gametes that have a mixture of maternal and paternal chromosomes. c. Fertilization is random, in that any two gametes can fuse.

**Exercise:**

**Problem:**

Compare the three main types of life cycles in multicellular organisms and give an example of an organism that employs each.

---

**Solution:**

a. In the haploid-dominant life cycle, the multicellular stage is haploid. The diploid stage is a spore that undergoes meiosis to produce cells that will divide mitotically to produce new multicellular organisms. Fungi have a haploid-dominant life cycle. b. In the diploid-dominant life cycle, the most visible or largest multicellular stage is diploid. The haploid stage is usually reduced to a single cell type, such as a gamete or spore. Animals, such as humans, have a diploid-dominant life cycle. c. In the alternation of generations life cycle, there are both haploid and diploid multicellular stages, although the haploid stage may be completely retained by the diploid stage. Plants have a life cycle with alternation of generations.

**Glossary**

alternation of generations

life-cycle type in which the diploid and haploid stages alternate

diploid-dominant

life-cycle type in which the multicellular diploid stage is prevalent

haploid-dominant

life-cycle type in which the multicellular haploid stage is prevalent

gametophyte

a multicellular haploid life-cycle stage that produces gametes

germ cells

specialized cell line that produces gametes, such as eggs or sperm

life cycle

the sequence of events in the development of an organism and the production of cells that produce offspring

sporophyte

a multicellular diploid life-cycle stage that produces haploid spores by meiosis

## Introduction

class="introduction"

Experimentin  
g with  
thousands of  
garden peas,  
Mendel  
uncovered the  
fundamentals  
of genetics.  
(credit:  
modification  
of work by  
Jerry  
Kirkhart)



Genetics is the study of heredity. Johann Gregor Mendel set the framework for genetics long before chromosomes or genes had been identified, at a time when meiosis was not well understood. Mendel selected a simple biological system and conducted methodical, quantitative analyses using



large sample sizes. Because of Mendel's work, the fundamental principles of heredity were revealed. We now know that genes, carried on chromosomes, are the basic functional units of heredity with the capability to be replicated, expressed, or mutated. Today, the postulates put forth by Mendel form the basis of classical, or Mendelian, genetics. Not all genes are transmitted from parents to offspring according to Mendelian genetics, but Mendel's experiments serve as an excellent starting point for thinking about inheritance.

## Mendel's Experiments and the Laws of Probability

By the end of this section, you will be able to:

- Describe the scientific reasons for the success of Mendel's experimental work
- Describe the expected outcomes of monohybrid crosses involving dominant and recessive alleles
- Apply the sum and product rules to calculate probabilities



Johann Gregor Mendel is considered the father of genetics.

Johann Gregor Mendel (1822–1884) ([\[link\]](#)) was a lifelong learner, teacher, scientist, and man of faith. As a young adult, he joined the Augustinian Abbey of St. Thomas in Brno in what is now the Czech Republic. Supported by the monastery, he taught physics, botany, and natural science courses at the secondary and university levels. In 1856, he began a decade-long research pursuit involving inheritance patterns in honeybees and plants, ultimately settling on pea plants as his primary **model system** (a system with convenient characteristics used to study a specific biological phenomenon to be applied to other systems). In 1865, Mendel presented the results of his experiments with nearly 30,000 pea plants to the local Natural History Society. He demonstrated that traits are transmitted faithfully from parents to offspring independently of other traits and in dominant and recessive patterns. In 1866, he published his work, *Experiments in Plant Hybridization*, [\[footnote\]](#) in the proceedings of the Natural History Society of Brünn.

Johann Gregor Mendel, *Versuche über Pflanzenhybriden Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr, 1865 Abhandlungen*, 3–47.

[for English translation see <http://www.mendelweb.org/Mendel.plain.html>]

Mendel's work went virtually unnoticed by the scientific community that believed, incorrectly, that the process of inheritance involved a blending of parental traits that produced an intermediate physical appearance in offspring; this hypothetical process appeared to be correct because of what we know now as continuous variation.

**Continuous variation** results from the action of many genes to determine a characteristic like human height. Offspring appear to be a “blend” of their parents' traits when we look at characteristics that exhibit continuous variation. The **blending theory of inheritance** asserted that the original parental traits were lost or absorbed by the blending in the offspring, but we now know that this is not the case. Mendel was the first researcher to see it. Instead of continuous characteristics, Mendel worked with traits that were inherited in distinct classes (specifically, violet versus white flowers); this is referred to as **discontinuous variation**. Mendel's choice of these kinds of traits allowed him to see experimentally that the traits were not blended in the offspring, nor were they absorbed, but rather that they kept their distinctness and could be passed on. In 1868, Mendel became abbot of the monastery and exchanged his scientific pursuits for his pastoral duties. He was not recognized for his extraordinary scientific contributions during his lifetime. In fact, it was not until 1900 that his work was rediscovered, reproduced, and revitalized by scientists on the brink of discovering the chromosomal basis of heredity.

## Mendel's Model System

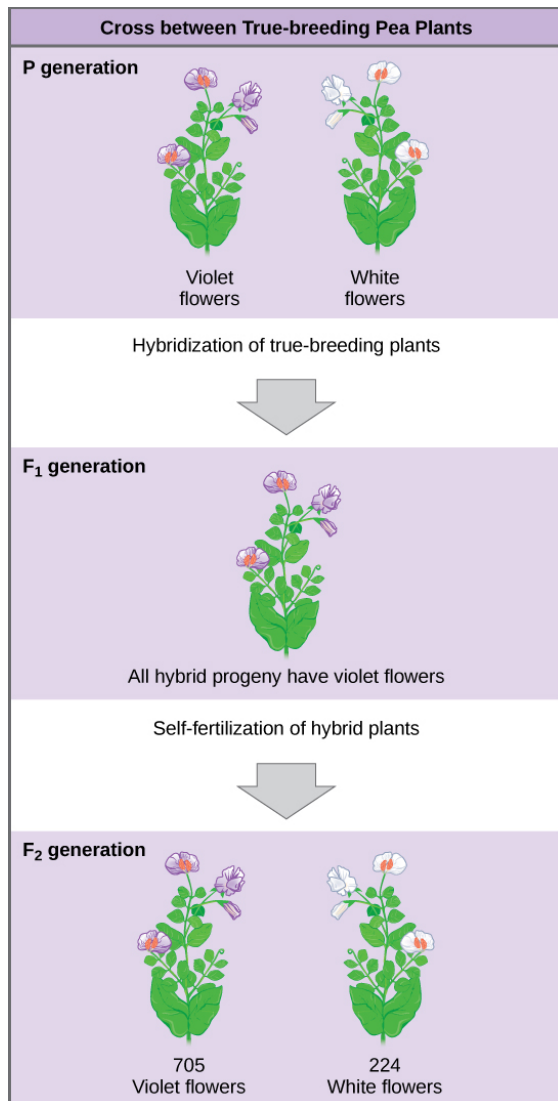
Mendel's seminal work was accomplished using the garden pea, *Pisum sativum*, to study inheritance. This species naturally self-fertilizes, such that pollen encounters ova within individual flowers. The flower petals remain sealed tightly until after pollination, preventing pollination from other plants. The result is highly inbred, or “true-breeding,” pea plants. These are plants that always produce offspring that look like the parent. By experimenting with true-breeding pea plants, Mendel avoided the appearance of unexpected traits in offspring that might occur if the plants were not true breeding. The garden pea also grows to maturity within one season, meaning that several generations could be evaluated over a relatively short time. Finally, large quantities of garden peas could be cultivated simultaneously, allowing Mendel to conclude that his results did not come about simply by chance.

## Mendelian Crosses

Mendel performed **hybridizations**, which involve mating two true-breeding individuals that have different traits. In the pea, which is naturally self-pollinating, this is done by manually transferring pollen from the anther of a mature pea plant of one variety to the stigma of a separate mature pea plant of the second variety. In plants, pollen carries the

male gametes (sperm) to the stigma, a sticky organ that traps pollen and allows the sperm to move down the pistil to the female gametes (ova) below. To prevent the pea plant that was receiving pollen from self-fertilizing and confounding his results, Mendel painstakingly removed all of the anthers from the plant's flowers before they had a chance to mature.

Plants used in first-generation crosses were called **P<sub>0</sub>**, or parental generation one, plants ([link](#)). Mendel collected the seeds belonging to the P<sub>0</sub> plants that resulted from each cross and grew them the following season. These offspring were called the **F<sub>1</sub>**, or the first filial (*filial* = offspring, daughter or son), generation. Once Mendel examined the characteristics in the F<sub>1</sub> generation of plants, he allowed them to self-fertilize naturally. He then collected and grew the seeds from the F<sub>1</sub> plants to produce the **F<sub>2</sub>**, or second filial, generation. Mendel's experiments extended beyond the F<sub>2</sub> generation to the F<sub>3</sub> and F<sub>4</sub> generations, and so on, but it was the ratio of characteristics in the P<sub>0</sub>-F<sub>1</sub>-F<sub>2</sub> generations that were the most intriguing and became the basis for Mendel's postulates.



In one of his experiments on inheritance patterns, Mendel crossed plants that were true-breeding for violet flower color with plants true-breeding for white flower color (the P generation).

The resulting hybrids in the F<sub>1</sub> generation all had violet flowers.

In the F<sub>2</sub> generation, approximately three quarters of the plants had violet flowers, and one quarter had white flowers.

## Garden Pea Characteristics Revealed the Basics of Heredity

In his 1865 publication, Mendel reported the results of his crosses involving seven different characteristics, each with two contrasting traits. A **trait** is defined as a variation in the physical appearance of a heritable characteristic. The characteristics included plant height, seed texture, seed color, flower color, pea pod size, pea pod color, and flower position. For the characteristic of flower color, for example, the two contrasting traits were white versus violet. To fully examine each characteristic, Mendel generated large numbers of  $F_1$  and  $F_2$  plants, reporting results from 19,959  $F_2$  plants alone. His findings were consistent.

What results did Mendel find in his crosses for flower color? First, Mendel confirmed that he had plants that bred true for white or violet flower color. Regardless of how many generations Mendel examined, all self-crossed offspring of parents with white flowers had white flowers, and all self-crossed offspring of parents with violet flowers had violet flowers. In addition, Mendel confirmed that, other than flower color, the pea plants were physically identical.

Once these validations were complete, Mendel applied the pollen from a plant with violet flowers to the stigma of a plant with white flowers. After gathering and sowing the seeds that resulted from this cross, Mendel found that 100 percent of the  $F_1$  hybrid generation had violet flowers. Conventional wisdom at that time would have predicted the hybrid flowers to be pale violet or for hybrid plants to have equal numbers of white and violet flowers. In other words, the contrasting parental traits were expected to blend in the offspring. Instead, Mendel's results demonstrated that the white flower trait in the  $F_1$  generation had completely disappeared.

Importantly, Mendel did not stop his experimentation there. He allowed the  $F_1$  plants to self-fertilize and found that, of  $F_2$ -generation plants, 705 had violet flowers and 224 had white flowers. This was a ratio of 3.15 violet flowers per one white flower, or approximately 3:1. When Mendel transferred pollen from a plant with violet flowers to the stigma of a plant with white flowers and vice versa, he obtained about the same ratio regardless of which parent, male or female, contributed which trait. This is called a **reciprocal cross**—a paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross. For the other six characteristics Mendel examined, the  $F_1$  and  $F_2$  generations behaved in the same way as they had for flower color. One of the two traits would disappear completely from the  $F_1$  generation only to reappear in the  $F_2$  generation at a ratio of approximately 3:1 ([link](#)).

The Results of Mendel's Garden Pea Hybridizations				
Characteristic	Contrasting P <sub>0</sub> Traits	F <sub>1</sub> Offspring Traits	F <sub>2</sub> Offspring Traits	F <sub>2</sub> Trait Ratios
Flower color	Violet vs. white	100 percent violet	705 violet 224 white	3.15:1
Flower position	Axial vs. terminal	100 percent axial	651 axial 207 terminal	3.14:1
Plant height	Tall vs. dwarf	100 percent tall	787 tall 277 dwarf	2.84:1
Seed texture	Round vs. wrinkled	100 percent round	5,474 round 1,850 wrinkled	2.96:1
Seed color	Yellow vs. green	100 percent yellow	6,022 yellow 2,001 green	3.01:1
Pea pod texture	Inflated vs. constricted	100 percent inflated	882 inflated 299 constricted	2.95:1

The Results of Mendel's Garden Pea Hybridizations				
Characteristic	Contrasting P <sub>0</sub> Traits	F <sub>1</sub> Offspring Traits	F <sub>2</sub> Offspring Traits	F <sub>2</sub> Trait Ratios
Pea pod color	Green vs. yellow	100 percent green	428 green 152 yellow	2.82:1

Upon compiling his results for many thousands of plants, Mendel concluded that the characteristics could be divided into expressed and latent traits. He called these, respectively, dominant and recessive traits. **Dominant traits** are those that are inherited unchanged in a hybridization. **Recessive traits** become latent, or disappear, in the offspring of a hybridization. The recessive trait does, however, reappear in the progeny of the hybrid offspring. An example of a dominant trait is the violet-flower trait. For this same characteristic (flower color), white-colored flowers are a recessive trait. The fact that the recessive trait reappeared in the F<sub>2</sub> generation meant that the traits remained separate (not blended) in the plants of the F<sub>1</sub> generation. Mendel also proposed that plants possessed two copies of the trait for the flower-color characteristic, and that each parent transmitted one of its two copies to its offspring, where they came together. Moreover, the physical observation of a dominant trait could mean that the genetic composition of the organism included two dominant versions of the characteristic or that it included one dominant and one recessive version. Conversely, the observation of a recessive trait meant that the organism lacked any dominant versions of this characteristic.

So why did Mendel repeatedly obtain 3:1 ratios in his crosses? To understand how Mendel deduced the basic mechanisms of inheritance that lead to such ratios, we must first review the laws of probability.

## Probability Basics

Probabilities are mathematical measures of likelihood. The empirical probability of an event is calculated by dividing the number of times the event occurs by the total number of opportunities for the event to occur. It is also possible to calculate theoretical probabilities by dividing the number of times that an event is expected to occur by the number of times that it could occur. Empirical probabilities come from observations, like those of Mendel. Theoretical probabilities come from knowing how the events are produced and assuming that the probabilities of individual outcomes are equal. A



probability of one for some event indicates that it is guaranteed to occur, whereas a probability of zero indicates that it is guaranteed not to occur. An example of a genetic event is a round seed produced by a pea plant. In his experiment, Mendel demonstrated that the probability of the event “round seed” occurring was one in the  $F_1$  offspring of true-breeding parents, one of which has round seeds and one of which has wrinkled seeds. When the  $F_1$  plants were subsequently self-crossed, the probability of any given  $F_2$  offspring having round seeds was now three out of four. In other words, in a large population of  $F_2$  offspring chosen at random, 75 percent were expected to have round seeds, whereas 25 percent were expected to have wrinkled seeds. Using large numbers of crosses, Mendel was able to calculate probabilities and use these to predict the outcomes of other crosses.

**The Product Rule and Sum Rule**

Mendel demonstrated that the pea-plant characteristics he studied were transmitted as discrete units from parent to offspring. As will be discussed, Mendel also determined that different characteristics, like seed color and seed texture, were transmitted independently of one another and could be considered in separate probability analyses. For instance, performing a cross between a plant with green, wrinkled seeds and a plant with yellow, round seeds still produced offspring that had a 3:1 ratio of green:yellow seeds (ignoring seed texture) and a 3:1 ratio of round:wrinkled seeds (ignoring seed color). The characteristics of color and texture did not influence each other.

The **product rule** of probability can be applied to this phenomenon of the independent transmission of characteristics. The product rule states that the probability of two independent events occurring together can be calculated by multiplying the individual probabilities of each event occurring alone. To demonstrate the product rule, imagine that you are rolling a six-sided die (D) and flipping a penny (P) at the same time. The die may roll any number from 1–6 ( $D_{\#}$ ), whereas the penny may turn up heads ( $P_H$ ) or tails ( $P_T$ ). The outcome of rolling the die has no effect on the outcome of flipping the penny and vice versa. There are 12 possible outcomes of this action ([\[link\]](#)), and each event is expected to occur with equal probability.

Twelve Equally Likely Outcomes of Rolling a Die and Flipping a Penny	
Rolling Die	Flipping Penny

Twelve Equally Likely Outcomes of Rolling a Die and Flipping a Penny	
Rolling Die	Flipping Penny
D <sub>1</sub>	P <sub>H</sub>
D <sub>1</sub>	P <sub>T</sub>
D <sub>2</sub>	P <sub>H</sub>
D <sub>2</sub>	P <sub>T</sub>
D <sub>3</sub>	P <sub>H</sub>
D <sub>3</sub>	P <sub>T</sub>
D <sub>4</sub>	P <sub>H</sub>
D <sub>4</sub>	P <sub>T</sub>
D <sub>5</sub>	P <sub>H</sub>
D <sub>5</sub>	P <sub>T</sub>
D <sub>6</sub>	P <sub>H</sub>
D <sub>6</sub>	P <sub>T</sub>

Of the 12 possible outcomes, the die has a 2/12 (or 1/6) probability of rolling a two, and the penny has a 6/12 (or 1/2) probability of coming up heads. By the product rule, the probability that you will obtain the combined outcome 2 and heads is: (D<sub>2</sub>) x (P<sub>H</sub>) = (1/6) x (1/2) or 1/12 ([link](#)). Notice the word “and” in the description of the probability. The “and” is a signal to apply the product rule. For example, consider how the product rule is applied to the dihybrid cross: the probability of having both dominant traits in the F<sub>2</sub> progeny is the product of the probabilities of having the dominant trait for each characteristic, as shown here:

**Equation:**

$$\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$$

On the other hand, the **sum rule** of probability is applied when considering two mutually exclusive outcomes that can come about by more than one pathway. The sum rule states that the probability of the occurrence of one event or the other event, of two mutually exclusive events, is the sum of their individual probabilities. Notice the word “or” in the description of the probability. The “or” indicates that you should apply the sum rule. In this case, let’s imagine you are flipping a penny (P) and a quarter (Q). What is the probability of one coin coming up heads and one coin coming up tails? This outcome can be achieved by two cases: the penny may be heads (P<sub>H</sub>) and the quarter may be tails (Q<sub>T</sub>), or the quarter may be heads (Q<sub>H</sub>) and the penny may be tails (P<sub>T</sub>). Either case fulfills the outcome. By the sum rule, we calculate the probability of obtaining one head and one tail as [(P<sub>H</sub>) × (Q<sub>T</sub>)] + [(Q<sub>H</sub>) × (P<sub>T</sub>)] = [(1/2) × (1/2)] + [(1/2) × (1/2)] = 1/2 ([link](#)). You should also notice that we used the product rule to calculate the probability of P<sub>H</sub> and Q<sub>T</sub>, and also the probability of P<sub>T</sub> and Q<sub>H</sub>, before we summed them. Again, the sum rule can be applied to show the probability of having just one dominant trait in the F<sub>2</sub> generation of a dihybrid cross:

**Equation:**

$$\frac{3}{16} + \frac{3}{4} = \frac{15}{16}$$

The Product Rule and Sum Rule	
Product Rule	Sum Rule
For independent events A and B, the probability (P) of them both occurring ( <i>A and B</i> ) is (P <sub>A</sub> × P <sub>B</sub> )	For mutually exclusive events A and B, the probability (P) that at least one occurs ( <i>A or B</i> ) is (P <sub>A</sub> + P <sub>B</sub> )

To use probability laws in practice, it is necessary to work with large sample sizes because small sample sizes are prone to deviations caused by chance. The large quantities of pea plants that Mendel examined allowed him calculate the probabilities of the traits appearing in his F<sub>2</sub> generation. As you will learn, this discovery meant that when parental traits were known, the offspring’s traits could be predicted accurately even before fertilization.

### Section Summary

Working with garden pea plants, Mendel found that crosses between parents that differed by one trait produced  $F_1$  offspring that all expressed the traits of one parent. Observable traits are referred to as dominant, and non-expressed traits are described as recessive. When the offspring in Mendel's experiment were self-crossed, the  $F_2$  offspring exhibited the dominant trait or the recessive trait in a 3:1 ratio, confirming that the recessive trait had been transmitted faithfully from the original  $P_0$  parent. Reciprocal crosses generated identical  $F_1$  and  $F_2$  offspring ratios. By examining sample sizes, Mendel showed that his crosses behaved reproducibly according to the laws of probability, and that the traits were inherited as independent events.

Two rules in probability can be used to find the expected proportions of offspring of different traits from different crosses. To find the probability of two or more independent events occurring together, apply the product rule and multiply the probabilities of the individual events. The use of the word “and” suggests the appropriate application of the product rule. To find the probability of two or more events occurring in combination, apply the sum rule and add their individual probabilities together. The use of the word “or” suggests the appropriate application of the sum rule.

## Review Questions

### Exercise:

#### Problem:

Mendel performed hybridizations by transferring pollen from the \_\_\_\_\_ of the male plant to the female ova.

- a. anther
- b. pistil
- c. stigma
- d. seed

---

#### Solution:

A

### Exercise:

#### Problem:

Which is one of the seven characteristics that Mendel observed in pea plants?

- a. flower size

- b. seed texture
- c. leaf shape
- d. stem color

---

**Solution:**

B

**Exercise:**

**Problem:**

Imagine you are performing a cross involving seed color in garden pea plants. What  $F_1$  offspring would you expect if you cross true-breeding parents with green seeds and yellow seeds? Yellow seed color is dominant over green.

- a. 100 percent yellow-green seeds
- b. 100 percent yellow seeds
- c. 50 percent yellow, 50 percent green seeds
- d. 25 percent green, 75 percent yellow seeds

---

**Solution:**

B

**Exercise:**

**Problem:**

Consider a cross to investigate the pea pod texture trait, involving constricted or inflated pods. Mendel found that the traits behave according to a dominant/recessive pattern in which inflated pods were dominant. If you performed this cross and obtained 650 inflated-pod plants in the  $F_2$  generation, approximately how many constricted-pod plants would you expect to have?

- a. 600
- b. 165
- c. 217
- d. 468

---

**Solution:**

C

## Free Response

### Exercise:

#### Problem:

Describe one of the reasons why the garden pea was an excellent choice of model system for studying inheritance.

---

#### Solution:

The garden pea is sessile and has flowers that close tightly during self-pollination. These features help to prevent accidental or unintentional fertilizations that could have diminished the accuracy of Mendel's data.

### Exercise:

#### Problem:

How would you perform a reciprocal cross for the characteristic of stem height in the garden pea?

---

#### Solution:

Two sets of  $P_0$  parents would be used. In the first cross, pollen would be transferred from a true-breeding tall plant to the stigma of a true-breeding dwarf plant. In the second cross, pollen would be transferred from a true-breeding dwarf plant to the stigma of a true-breeding tall plant. For each cross,  $F_1$  and  $F_2$  offspring would be analyzed to determine if offspring traits were affected according to which parent donated each trait.

## Glossary

### blending theory of inheritance

hypothetical inheritance pattern in which parental traits are blended together in the offspring to produce an intermediate physical appearance

### continuous variation

inheritance pattern in which a character shows a range of trait values with small gradations rather than large gaps between them

### discontinuous variation

inheritance pattern in which traits are distinct and are transmitted independently of one another

dominant

trait which confers the same physical appearance whether an individual has two copies of the trait or one copy of the dominant trait and one copy of the recessive trait

$F_1$

first filial generation in a cross; the offspring of the parental generation

$F_2$

second filial generation produced when  $F_1$  individuals are self-crossed or fertilized with each other

hybridization

process of mating two individuals that differ with the goal of achieving a certain characteristic in their offspring

model system

species or biological system used to study a specific biological phenomenon to be applied to other different species

$P_0$

parental generation in a cross

product rule

probability of two independent events occurring simultaneously can be calculated by multiplying the individual probabilities of each event occurring alone

recessive

trait that appears “latent” or non-expressed when the individual also carries a dominant trait for that same characteristic; when present as two identical copies, the recessive trait is expressed

reciprocal cross

paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross

sum rule

probability of the occurrence of at least one of two mutually exclusive events is the sum of their individual probabilities

trait

variation in the physical appearance of a heritable characteristic

## Characteristics and Traits

By the end of this section, you will be able to:

- Explain the relationship between genotypes and phenotypes in dominant and recessive gene systems
- Develop a Punnett square to calculate the expected proportions of genotypes and phenotypes in a monohybrid cross
- Explain the purpose and methods of a test cross
- Identify non-Mendelian inheritance patterns such as incomplete dominance, codominance, recessive lethals, multiple alleles, and sex linkage

The seven characteristics that Mendel evaluated in his pea plants were each expressed as one of two versions, or traits. The physical expression of characteristics is accomplished through the expression of genes carried on chromosomes. The genetic makeup of peas consists of two similar or homologous copies of each chromosome, one from each parent. Each pair of homologous chromosomes has the same linear order of genes. In other words, peas are diploid organisms in that they have two copies of each chromosome. The same is true for many other plants and for virtually all animals. Diploid organisms utilize meiosis to produce haploid gametes, which contain one copy of each homologous chromosome that unite at fertilization to create a diploid zygote.

For cases in which a single gene controls a single characteristic, a diploid organism has two genetic copies that may or may not encode the same version of that characteristic. Gene variants that arise by mutation and exist at the same relative locations on homologous chromosomes are called **alleles**. Mendel examined the inheritance of genes with just two allele forms, but it is common to encounter more than two alleles for any given gene in a natural population.

## Phenotypes and Genotypes

Two alleles for a given gene in a diploid organism are expressed and interact to produce physical characteristics. The observable traits expressed by an organism are referred to as its **phenotype**. An organism's underlying



genetic makeup, consisting of both physically visible and non-expressed alleles, is called its **genotype**. Mendel's hybridization experiments demonstrate the difference between phenotype and genotype. When true-breeding plants in which one parent had yellow pods and one had green pods were cross-fertilized, all of the  $F_1$  hybrid offspring had yellow pods. That is, the hybrid offspring were phenotypically identical to the true-breeding parent with yellow pods. However, we know that the allele donated by the parent with green pods was not simply lost because it reappeared in some of the  $F_2$  offspring. Therefore, the  $F_1$  plants must have been genotypically different from the parent with yellow pods.

The  $P_1$  plants that Mendel used in his experiments were each homozygous for the trait he was studying. Diploid organisms that are **homozygous** at a given gene, or locus, have two identical alleles for that gene on their homologous chromosomes. Mendel's parental pea plants always bred true because both of the gametes produced carried the same trait. When  $P_1$  plants with contrasting traits were cross-fertilized, all of the offspring were **heterozygous** for the contrasting trait, meaning that their genotype reflected that they had different alleles for the gene being examined.

## Dominant and Recessive Alleles

Our discussion of homozygous and heterozygous organisms brings us to why the  $F_1$  heterozygous offspring were identical to one of the parents, rather than expressing both alleles. In all seven pea-plant characteristics, one of the two contrasting alleles was dominant, and the other was recessive. Mendel called the dominant allele the expressed unit factor; the recessive allele was referred to as the latent unit factor. We now know that these so-called unit factors are actually genes on homologous chromosome pairs. For a gene that is expressed in a dominant and recessive pattern, homozygous dominant and heterozygous organisms will look identical (that is, they will have different genotypes but the same phenotype). The recessive allele will only be observed in homozygous recessive individuals ([link](#)).

Human Inheritance in Dominant and Recessive Patterns	
Dominant Traits	Recessive Traits
Achondroplasia	Albinism
Brachydactyly	Cystic fibrosis
Huntington's disease	Duchenne muscular dystrophy
Marfan syndrome	Galactosemia
Neurofibromatosis	Phenylketonuria
Widow's peak	Sickle-cell anemia
Wooly hair	Tay-Sachs disease

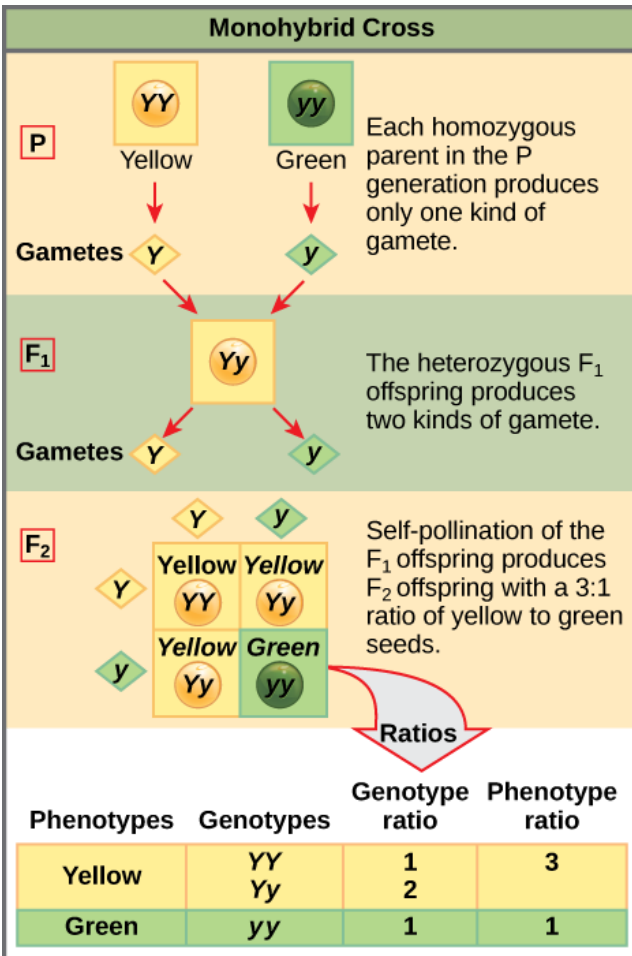
Several conventions exist for referring to genes and alleles. For the purposes of this chapter, we will abbreviate genes using the first letter of the gene's corresponding dominant trait. For example, violet is the dominant trait for a pea plant's flower color, so the flower-color gene would be abbreviated as *V* (note that it is customary to italicize gene designations). Furthermore, we will use uppercase and lowercase letters to represent dominant and recessive alleles, respectively. Therefore, we would refer to the genotype of a homozygous dominant pea plant with violet flowers as *VV*, a homozygous recessive pea plant with white flowers as *vv*, and a heterozygous pea plant with violet flowers as *Vv*.

## The Punnett Square Approach for a Monohybrid Cross

When fertilization occurs between two true-breeding parents that differ in only one characteristic, the process is called a **monohybrid** cross, and the resulting offspring are monohybrids. Mendel performed seven monohybrid crosses involving contrasting traits for each characteristic. On the basis of his results in  $F_1$  and  $F_2$  generations, Mendel postulated that each parent in

the monohybrid cross contributed one of two paired unit factors to each offspring, and every possible combination of unit factors was equally likely.

To demonstrate a monohybrid cross, consider the case of true-breeding pea plants with yellow versus green pea seeds. The dominant seed color is yellow; therefore, the parental genotypes were YY for the plants with yellow seeds and yy for the plants with green seeds, respectively. A **Punnett square**, devised by the British geneticist Reginald Punnett, can be drawn that applies the rules of probability to predict the possible outcomes of a genetic cross or mating and their expected frequencies. To prepare a Punnett square, all possible combinations of the parental alleles are listed along the top (for one parent) and side (for the other parent) of a grid, representing their meiotic segregation into haploid gametes. Then the combinations of egg and sperm are made in the boxes in the table to show which alleles are combining. Each box then represents the diploid genotype of a zygote, or fertilized egg, that could result from this mating. Because each possibility is equally likely, genotypic ratios can be determined from a Punnett square. If the pattern of inheritance (dominant or recessive) is known, the phenotypic ratios can be inferred as well. For a monohybrid cross of two true-breeding parents, each parent contributes one type of allele. In this case, only one genotype is possible. All offspring are Yy and have yellow seeds ([\[link\]](#)).



In the P generation, pea plants that are true-breeding for the dominant yellow phenotype are crossed with plants with the recessive green phenotype. This cross produces F<sub>1</sub> heterozygotes with a yellow phenotype. Punnett square analysis can be used to predict the genotypes of the F<sub>2</sub> generation.

A self-cross of one of the Yy heterozygous offspring can be represented in a 2 × 2 Punnett square because each parent can donate one of two different alleles. Therefore, the offspring can potentially have one of four allele

combinations: YY, Yy, yY, or yy ([link](#)). Notice that there are two ways to obtain the Yy genotype: a Y from the egg and a y from the sperm, or a y from the egg and a Y from the sperm. Both of these possibilities must be counted. Recall that Mendel's pea-plant characteristics behaved in the same way in reciprocal crosses. Therefore, the two possible heterozygous combinations produce offspring that are genotypically and phenotypically identical despite their dominant and recessive alleles deriving from different parents. They are grouped together. Because fertilization is a random event, we expect each combination to be equally likely and for the offspring to exhibit a ratio of YY:Yy:yy genotypes of 1:2:1 ([link](#)). Furthermore, because the YY and Yy offspring have yellow seeds and are phenotypically identical, applying the sum rule of probability, we expect the offspring to exhibit a phenotypic ratio of 3 yellow:1 green. Indeed, working with large sample sizes, Mendel observed approximately this ratio in every F<sub>2</sub> generation resulting from crosses for individual traits.

Mendel validated these results by performing an F<sub>3</sub> cross in which he self-crossed the dominant- and recessive-expressing F<sub>2</sub> plants. When he self-crossed the plants expressing green seeds, all of the offspring had green seeds, confirming that all green seeds had homozygous genotypes of yy. When he self-crossed the F<sub>2</sub> plants expressing yellow seeds, he found that one-third of the plants bred true, and two-thirds of the plants segregated at a 3:1 ratio of yellow:green seeds. In this case, the true-breeding plants had homozygous (YY) genotypes, whereas the segregating plants corresponded to the heterozygous (Yy) genotype. When these plants self-fertilized, the outcome was just like the F<sub>1</sub> self-fertilizing cross.

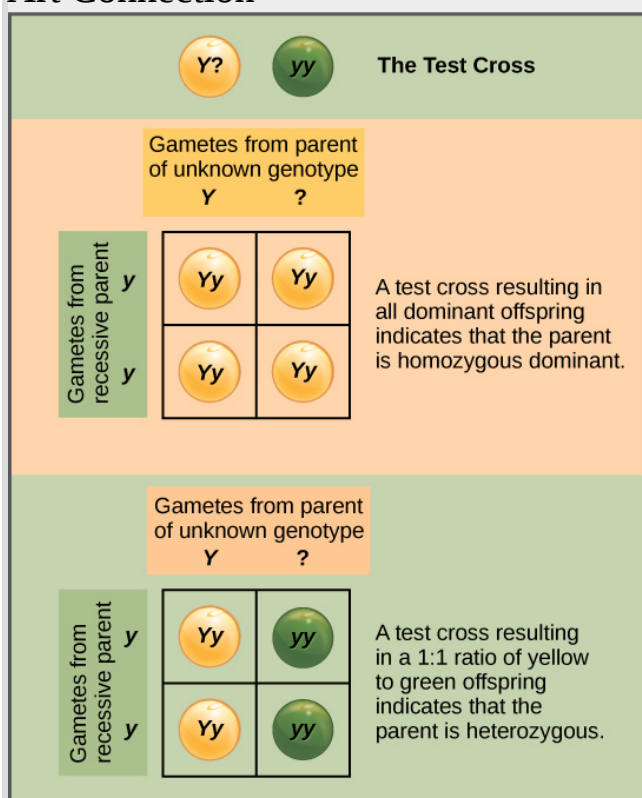
## The Test Cross Distinguishes the Dominant Phenotype

Beyond predicting the offspring of a cross between known homozygous or heterozygous parents, Mendel also developed a way to determine whether an organism that expressed a dominant trait was a heterozygote or a homozygote. Called the **test cross**, this technique is still used by plant and animal breeders. In a test cross, the dominant-expressing organism is crossed with an organism that is homozygous recessive for the same characteristic. If the dominant-expressing organism is a homozygote, then

all  $F_1$  offspring will be heterozygotes expressing the dominant trait ([\[link\]](#)). Alternatively, if the dominant expressing organism is a heterozygote, the  $F_1$  offspring will exhibit a 1:1 ratio of heterozygotes and recessive homozygotes ([\[link\]](#)). The test cross further validates Mendel's postulate that pairs of unit factors segregate equally.

### Note:

#### Art Connection



A test cross can be performed to determine whether an organism expressing a dominant trait is a homozygote or a heterozygote.

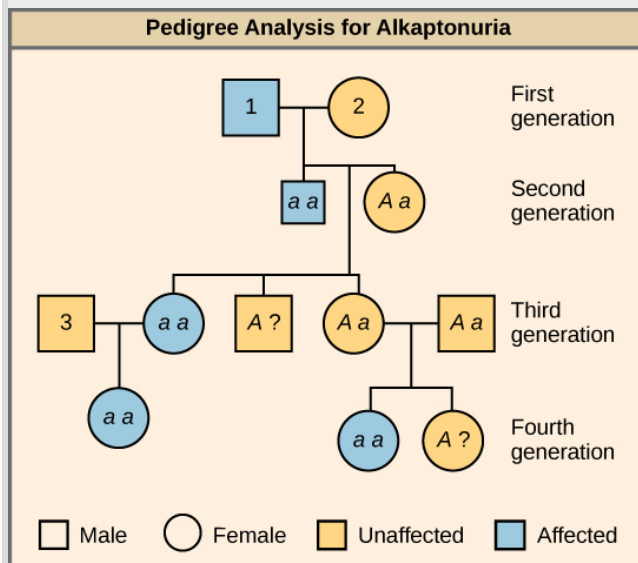
In pea plants, round peas ( $R$ ) are dominant to wrinkled peas ( $r$ ). You do a test cross between a pea plant with wrinkled peas (genotype  $rr$ ) and a plant of unknown genotype that has round peas. You end up with three plants, all

which have round peas. From this data, can you tell if the round pea parent plant is homozygous dominant or heterozygous? If the round pea parent plant is heterozygous, what is the probability that a random sample of 3 progeny peas will all be round?

Many human diseases are genetically inherited. A healthy person in a family in which some members suffer from a recessive genetic disorder may want to know if he or she has the disease-causing gene and what risk exists of passing the disorder on to his or her offspring. Of course, doing a test cross in humans is unethical and impractical. Instead, geneticists use **pedigree analysis** to study the inheritance pattern of human genetic diseases ([link](#)).

**Note:**

**Art Connection**



Alkaptonuria is a recessive genetic disorder in which two amino acids, phenylalanine and tyrosine, are not properly metabolized. Affected individuals may have

darkened skin and brown urine, and may suffer joint damage and other complications. In this pedigree, individuals with the disorder are indicated in blue and have the genotype  $aa$ . Unaffected individuals are indicated in yellow and have the genotype  $AA$  or  $Aa$ .

Note that it is often possible to determine a person's genotype from the genotype of their offspring. For example, if neither parent has the disorder but their child does, they must be heterozygous. Two individuals on the pedigree have an unaffected phenotype but unknown genotype.

Because they do not have the disorder, they must have at least one normal allele, so their genotype gets the “A?” designation.

What are the genotypes of the individuals labeled 1, 2 and 3?

## Alternatives to Dominance and Recessiveness

Mendel's experiments with pea plants suggested that: (1) two “units” or alleles exist for every gene; (2) alleles maintain their integrity in each generation (no blending); and (3) in the presence of the dominant allele, the recessive allele is hidden and makes no contribution to the phenotype. Therefore, recessive alleles can be “carried” and not expressed by individuals. Such heterozygous individuals are sometimes referred to as “carriers.” Further genetic studies in other plants and animals have shown



that much more complexity exists, but that the fundamental principles of Mendelian genetics still hold true. In the sections to follow, we consider some of the extensions of Mendelism. If Mendel had chosen an experimental system that exhibited these genetic complexities, it's possible that he would not have understood what his results meant.

## Incomplete Dominance

Mendel's results, that traits are inherited as dominant and recessive pairs, contradicted the view at that time that offspring exhibited a blend of their parents' traits. However, the heterozygote phenotype occasionally does appear to be intermediate between the two parents. For example, in the snapdragon, *Antirrhinum majus* ([\[link\]](#)), a cross between a homozygous parent with white flowers ( $C^W C^W$ ) and a homozygous parent with red flowers ( $C^R C^R$ ) will produce offspring with pink flowers ( $C^R C^W$ ). (Note that different genotypic abbreviations are used for Mendelian extensions to distinguish these patterns from simple dominance and recessiveness.) This pattern of inheritance is described as **incomplete dominance**, denoting the expression of two contrasting alleles such that the individual displays an intermediate phenotype. The allele for red flowers is incompletely dominant over the allele for white flowers. However, the results of a heterozygote self-cross can still be predicted, just as with Mendelian dominant and recessive crosses. In this case, the genotypic ratio would be 1  $C^R C^R$ :2  $C^R C^W$ :1  $C^W C^W$ , and the phenotypic ratio would be 1:2:1 for red:pink:white.



These pink flowers of a heterozygote snapdragon result from incomplete dominance.

(credit:  
“storebukkebruse”/Flickr)

## Codominance





A variation on incomplete dominance is **codominance**, in which both alleles for the same characteristic are simultaneously expressed in the heterozygote. An example of codominance is the MN blood groups of humans. The M and N alleles are expressed in the form of an M or N antigen present on the surface of red blood cells. Homozygotes ( $L^M L^M$  and  $L^N L^N$ ) express either the M or the N allele, and heterozygotes ( $L^M L^N$ ) express both alleles equally. In a self-cross between heterozygotes

expressing a codominant trait, the three possible offspring genotypes are phenotypically distinct. However, the 1:2:1 genotypic ratio characteristic of a Mendelian monohybrid cross still applies.

## Multiple Alleles

Mendel implied that only two alleles, one dominant and one recessive, could exist for a given gene. We now know that this is an oversimplification. Although individual humans (and all diploid organisms) can only have two alleles for a given gene, multiple alleles may exist at the population level such that many combinations of two alleles are observed. Note that when many alleles exist for the same gene, the convention is to denote the most common phenotype or genotype among wild animals as the **wild type** (often abbreviated “+”); this is considered the standard or norm. All other phenotypes or genotypes are considered **variants** of this standard, meaning that they deviate from the wild type. The variant may be recessive or dominant to the wild-type allele.

An example of multiple alleles is coat color in rabbits ([link](#)). Here, four alleles exist for the  $c$  gene. The wild-type version,  $C^+C^+$ , is expressed as brown fur. The chinchilla phenotype,  $c^{ch}c^{ch}$ , is expressed as black-tipped white fur. The Himalayan phenotype,  $c^hc^h$ , has black fur on the extremities and white fur elsewhere. Finally, the albino, or “colorless” phenotype,  $cc$ , is expressed as white fur. In cases of multiple alleles, dominance hierarchies can exist. In this case, the wild-type allele is dominant over all the others, chinchilla is incompletely dominant over Himalayan and albino, and Himalayan is dominant over albino. This hierarchy, or allelic series, was revealed by observing the phenotypes of each possible heterozygote offspring.

Allele			
C	c <sup>ch</sup>	c <sup>h</sup>	c
Genotype			
CC	c <sup>ch</sup> c <sup>ch</sup>	c <sup>h</sup> c <sup>h</sup>	cc
Phenotype			
WILD TYPE: Brown fur	CHINCHILLA: Black-tipped white fur	HIMALAYAN: White fur with black paws, nose, ears, tail	ALBINO: White fur
			

Four different alleles exist for the rabbit coat color (C) gene.

The complete dominance of a wild-type phenotype over all other mutants often occurs as an effect of “dosage” of a specific gene product, such that the wild-type allele supplies the correct amount of gene product whereas the mutant alleles cannot. For the allelic series in rabbits, the wild-type allele may supply a given dosage of fur pigment, whereas the mutants supply a lesser dosage or none at all. Interestingly, the Himalayan phenotype is the result of an allele that produces a temperature-sensitive gene product that only produces pigment in the cooler extremities of the rabbit’s body.

Alternatively, one mutant allele can be dominant over all other phenotypes, including the wild type. This may occur when the mutant allele somehow interferes with the genetic message so that even a heterozygote with one wild-type allele copy expresses the mutant phenotype. One way in which the mutant allele can interfere is by enhancing the function of the wild-type gene product or changing its distribution in the body. One example of this is the *Antennapedia* mutation in *Drosophila* ([link](#)). In this case, the mutant allele expands the distribution of the gene product, and as a result, the

*Antennapedia* heterozygote develops legs on its head where its antennae should be.



As seen in comparing the wild-type *Drosophila* (left) and the *Antennapedia* mutant (right), the *Antennapedia* mutant has legs on its head in place of antennae.

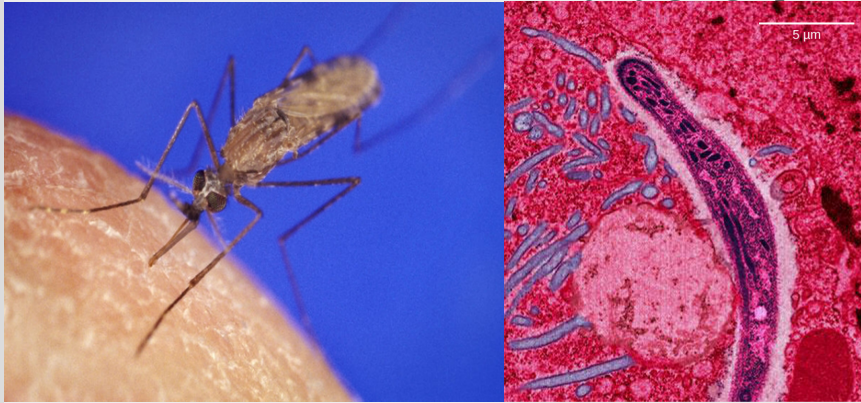
**Note:**

Evolution Connection

**Multiple Alleles Confer Drug Resistance in the Malaria Parasite**

Malaria is a parasitic disease in humans that is transmitted by infected female mosquitoes, including *Anopheles gambiae* ([\[link\]a](#)), and is characterized by cyclic high fevers, chills, flu-like symptoms, and severe anemia. *Plasmodium falciparum* and *P. vivax* are the most common causative agents of malaria, and *P. falciparum* is the most deadly ([\[link\]b](#)). When promptly and correctly treated, *P. falciparum* malaria has a mortality

rate of 0.1 percent. However, in some parts of the world, the parasite has evolved resistance to commonly used malaria treatments, so the most effective malarial treatments can vary by geographic region.



The (a) *Anopheles gambiae*, or African malaria mosquito, acts as a vector in the transmission to humans of the malaria-causing parasite (b) *Plasmodium falciparum*, here visualized using false-color transmission electron microscopy. (credit a: James D. Gathany; credit b: Ute Frevert; false color by Margaret Shear; scale-bar data from Matt Russell)

In Southeast Asia, Africa, and South America, *P. falciparum* has developed resistance to the anti-malarial drugs chloroquine, mefloquine, and sulfadoxine-pyrimethamine. *P. falciparum*, which is haploid during the life stage in which it is infectious to humans, has evolved multiple drug-resistant mutant alleles of the *dhps* gene. Varying degrees of sulfadoxine resistance are associated with each of these alleles. Being haploid, *P. falciparum* needs only one drug-resistant allele to express this trait. In Southeast Asia, different sulfadoxine-resistant alleles of the *dhps* gene are localized to different geographic regions. This is a common evolutionary phenomenon that occurs because drug-resistant mutants arise in a population and interbreed with other *P. falciparum* isolates in close proximity. Sulfadoxine-resistant parasites cause considerable human hardship in regions where this drug is widely used as an over-the-counter malaria remedy. As is common with pathogens that multiply to large



numbers within an infection cycle, *P. falciparum* evolves relatively rapidly (over a decade or so) in response to the selective pressure of commonly used anti-malarial drugs. For this reason, scientists must constantly work to develop new drugs or drug combinations to combat the worldwide malaria burden. [\[footnote\]](#)

Sumiti Vinayak, et al., “Origin and Evolution of Sulfadoxine Resistant *Plasmodium falciparum*,” *Public Library of Science Pathogens* 6, no. 3 (2010): e1000830, doi:10.1371/journal.ppat.1000830.

## X-Linked Traits

In humans, as well as in many other animals and some plants, the sex of the individual is determined by sex chromosomes. The sex chromosomes are one pair of non-homologous chromosomes. Until now, we have only considered inheritance patterns among non-sex chromosomes, or **autosomes**. In addition to 22 homologous pairs of autosomes, human females have a homologous pair of X chromosomes, whereas human males have an XY chromosome pair. Although the Y chromosome contains a small region of similarity to the X chromosome so that they can pair during meiosis, the Y chromosome is much shorter and contains many fewer genes. When a gene being examined is present on the X chromosome, but not on the Y chromosome, it is said to be **X-linked**.

Eye color in *Drosophila* was one of the first X-linked traits to be identified. Thomas Hunt Morgan mapped this trait to the X chromosome in 1910. Like humans, *Drosophila* males have an XY chromosome pair, and females are XX. In flies, the wild-type eye color is red ( $X^W$ ) and it is dominant to white eye color ( $X^w$ ) ([\[link\]](#)). Because of the location of the eye-color gene, reciprocal crosses do not produce the same offspring ratios. Males are said to be **hemizygous**, because they have only one allele for any X-linked characteristic. Hemizyosity makes the descriptions of dominance and recessiveness irrelevant for XY males. *Drosophila* males lack a second allele copy on the Y chromosome; that is, their genotype can only be  $X^WY$  or  $X^wY$ . In contrast, females have two allele copies of this gene and can be  $X^WX^W$ ,  $X^WX^w$ , or  $X^wX^w$ .



In *Drosophila*, several genes determine eye color. The genes for white and vermilion eye colors are located on the X chromosome. Others are located on the autosomes. Clockwise from top left are brown, cinnabar, sepia, vermilion, white, and red. Red eye color is wild-type and is dominant to white eye color.

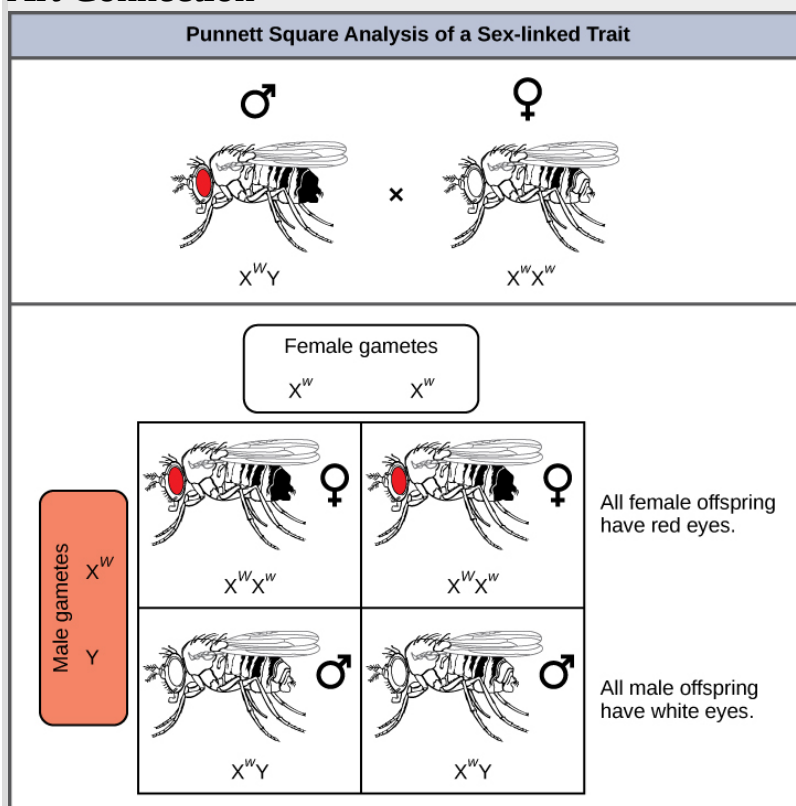
In an X-linked cross, the genotypes of  $F_1$  and  $F_2$  offspring depend on whether the recessive trait was expressed by the male or the female in the  $P_1$  generation. With regard to *Drosophila* eye color, when the  $P_1$  male expresses the white-eye phenotype and the female is homozygous red-eyed, all members of the  $F_1$  generation exhibit red eyes ([link](#)). The  $F_1$  females



are heterozygous ( $X^W X^w$ ), and the males are all  $X^W Y$ , having received their X chromosome from the homozygous dominant  $P_1$  female and their Y chromosome from the  $P_1$  male. A subsequent cross between the  $X^W X^w$  female and the  $X^W Y$  male would produce only red-eyed females (with  $X^W X^W$  or  $X^W X^w$  genotypes) and both red- and white-eyed males (with  $X^W Y$  or  $X^w Y$  genotypes). Now, consider a cross between a homozygous white-eyed female and a male with red eyes. The  $F_1$  generation would exhibit only heterozygous red-eyed females ( $X^W X^w$ ) and only white-eyed males ( $X^w Y$ ). Half of the  $F_2$  females would be red-eyed ( $X^W X^W$ ) and half would be white-eyed ( $X^w X^w$ ). Similarly, half of the  $F_2$  males would be red-eyed ( $X^W Y$ ) and half would be white-eyed ( $X^w Y$ ).

### Note:

### Art Connection



Punnett square analysis is used to determine the ratio of offspring from a cross between a

red-eyed male fruit fly and a white-eyed  
female fruit fly.

What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

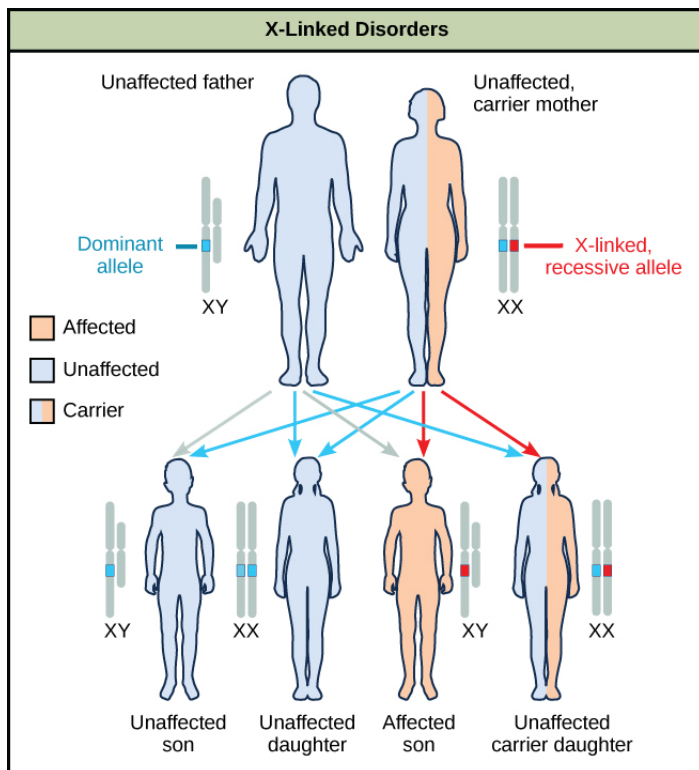
Discoveries in fruit fly genetics can be applied to human genetics. When a female parent is homozygous for a recessive X-linked trait, she will pass the trait on to 100 percent of her offspring. Her male offspring are, therefore, destined to express the trait, as they will inherit their father's Y chromosome. In humans, the alleles for certain conditions (some forms of color blindness, hemophilia, and muscular dystrophy) are X-linked. Females who are heterozygous for these diseases are said to be carriers and may not exhibit any phenotypic effects. These females will pass the disease to half of their sons and will pass carrier status to half of their daughters; therefore, recessive X-linked traits appear more frequently in males than females.

In some groups of organisms with sex chromosomes, the sex with the non-homologous sex chromosomes is the female rather than the male. This is the case for all birds. In this case, sex-linked traits will be more likely to appear in the female, in which they are hemizygous.

## **Human Sex-linked Disorders**

Sex-linkage studies in Morgan's laboratory provided the fundamentals for understanding X-linked recessive disorders in humans, which include red-green color blindness, and Types A and B hemophilia. Because human males need to inherit only one recessive mutant X allele to be affected, X-linked disorders are disproportionately observed in males. Females must inherit recessive X-linked alleles from both of their parents in order to express the trait. When they inherit one recessive X-linked mutant allele and one dominant X-linked wild-type allele, they are carriers of the trait and are typically unaffected. Carrier females can manifest mild forms of the trait

due to the inactivation of the dominant allele located on one of the X chromosomes. However, female carriers can contribute the trait to their sons, resulting in the son exhibiting the trait, or they can contribute the recessive allele to their daughters, resulting in the daughters being carriers of the trait ([link](#)). Although some Y-linked recessive disorders exist, typically they are associated with infertility in males and are therefore not transmitted to subsequent generations.



The son of a woman who is a carrier of a recessive X-linked disorder will have a 50 percent chance of being affected. A daughter will not be affected, but she will have a 50 percent chance of being a carrier like her mother.

**Note:**

## Link to Learning



Watch this video to learn more about sex-linked traits.

[https://www.openstaxcollege.org/l/sex-linked\\_traits](https://www.openstaxcollege.org/l/sex-linked_traits)

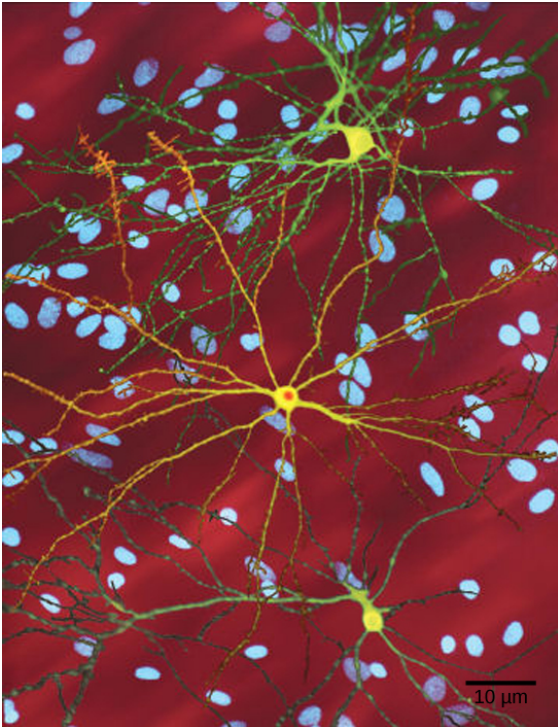
## Lethality

A large proportion of genes in an individual's genome are essential for survival. Occasionally, a nonfunctional allele for an essential gene can arise by mutation and be transmitted in a population as long as individuals with this allele also have a wild-type, functional copy. The wild-type allele functions at a capacity sufficient to sustain life and is therefore considered to be dominant over the nonfunctional allele. However, consider two heterozygous parents that have a genotype of wild-type/nonfunctional mutant for a hypothetical essential gene. In one quarter of their offspring, we would expect to observe individuals that are homozygous recessive for the nonfunctional allele. Because the gene is essential, these individuals might fail to develop past fertilization, die *in utero*, or die later in life, depending on what life stage requires this gene. An inheritance pattern in which an allele is only lethal in the homozygous form and in which the heterozygote may be normal or have some altered non-lethal phenotype is referred to as **recessive lethal**.

For crosses between heterozygous individuals with a recessive lethal allele that causes death before birth when homozygous, only wild-type homozygotes and heterozygotes would be observed. The genotypic ratio would therefore be 2:1. In other instances, the recessive lethal allele might also exhibit a dominant (but not lethal) phenotype in the heterozygote. For

instance, the recessive lethal *Curly* allele in *Drosophila* affects wing shape in the heterozygote form but is lethal in the homozygote.

A single copy of the wild-type allele is not always sufficient for normal functioning or even survival. The **dominant lethal** inheritance pattern is one in which an allele is lethal both in the homozygote and the heterozygote; this allele can only be transmitted if the lethality phenotype occurs after reproductive age. Individuals with mutations that result in dominant lethal alleles fail to survive even in the heterozygote form. Dominant lethal alleles are very rare because, as you might expect, the allele only lasts one generation and is not transmitted. However, just as the recessive lethal allele might not immediately manifest the phenotype of death, dominant lethal alleles also might not be expressed until adulthood. Once the individual reaches reproductive age, the allele may be unknowingly passed on, resulting in a delayed death in both generations. An example of this in humans is Huntington's disease, in which the nervous system gradually wastes away ([\[link\]](#)). People who are heterozygous for the dominant Huntington allele (*Hh*) will inevitably develop the fatal disease. However, the onset of Huntington's disease may not occur until age 40, at which point the afflicted persons may have already passed the allele to 50 percent of their offspring.



The neuron in the center of this micrograph (yellow) has nuclear inclusions characteristic of Huntington's disease (orange area in the center of the neuron).

Huntington's disease occurs when an abnormal dominant allele for the Huntington gene is present. (credit: Dr. Steven Finkbeiner, Gladstone Institute of Neurological Disease, The Taube-Koret Center for Huntington's Disease Research, and the University of California San Francisco/Wikimedia)

## Section Summary

When true-breeding or homozygous individuals that differ for a certain trait are crossed, all of the offspring will be heterozygotes for that trait. If the traits are inherited as dominant and recessive, the  $F_1$  offspring will all exhibit the same phenotype as the parent homozygous for the dominant trait. If these heterozygous offspring are self-crossed, the resulting  $F_2$  offspring will be equally likely to inherit gametes carrying the dominant or recessive trait, giving rise to offspring of which one quarter are homozygous dominant, half are heterozygous, and one quarter are homozygous recessive. Because homozygous dominant and heterozygous individuals are phenotypically identical, the observed traits in the  $F_2$  offspring will exhibit a ratio of three dominant to one recessive.

Alleles do not always behave in dominant and recessive patterns. Incomplete dominance describes situations in which the heterozygote exhibits a phenotype that is intermediate between the homozygous phenotypes. Codominance describes the simultaneous expression of both of the alleles in the heterozygote. Although diploid organisms can only have two alleles for any given gene, it is common for more than two alleles of a gene to exist in a population. In humans, as in many animals and some plants, females have two X chromosomes and males have one X and one Y chromosome. Genes that are present on the X but not the Y chromosome are said to be X-linked, such that males only inherit one allele for the gene, and females inherit two. Finally, some alleles can be lethal. Recessive lethal alleles are only lethal in homozygotes, but dominant lethal alleles are fatal in heterozygotes as well.

## Art Connections

### Exercise:

**Problem:**

[\[link\]](#) In pea plants, round peas ( $R$ ) are dominant to wrinkled peas ( $r$ ). You do a test cross between a pea plant with wrinkled peas (genotype  $rr$ ) and a plant of unknown genotype that has round peas. You end up with three plants, all which have round peas. From this data, can you tell if the round pea parent plant is homozygous dominant or heterozygous? If the round pea parent plant is heterozygous, what is the probability that a random sample of 3 progeny peas will all be round?

---

**Solution:**

[\[link\]](#) You cannot be sure if the plant is homozygous or heterozygous as the data set is too small: by random chance, all three plants might have acquired only the dominant gene even if the recessive one is present. If the round pea parent is heterozygous, there is a one-eighth probability that a random sample of three progeny peas will all be round.

**Exercise:****Problem:**

[\[link\]](#) What are the genotypes of the individuals labeled 1, 2 and 3?

---

**Solution:**

[\[link\]](#) Individual 1 has the genotype  $aa$ . Individual 2 has the genotype  $Aa$ . Individual 3 has the genotype  $Aa$ .

**Exercise:****Problem:**

[\[link\]](#) What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

---

**Solution:**



[\[link\]](#) Half of the female offspring would be heterozygous ( $X^W X^w$ ) with red eyes, and half would be homozygous recessive ( $X^w X^w$ ) with white eyes. Half of the male offspring would be hemizygous dominant ( $X^W Y$ ) with red eyes, and half would be hemizygous recessive ( $X^w Y$ ) with white eyes.

## Review Questions

### Exercise:

#### Problem:

The observable traits expressed by an organism are described as its \_\_\_\_\_.

- a. phenotype
- b. genotype
- c. alleles
- d. zygote

---

#### Solution:

A

### Exercise:

#### Problem:

A recessive trait will be observed in individuals that are \_\_\_\_\_ for that trait.

- a. heterozygous
- b. homozygous or heterozygous
- c. homozygous
- d. diploid

---

#### Solution:

C

**Exercise:**

**Problem:**

If black and white true-breeding mice are mated and the result is all gray offspring, what inheritance pattern would this be indicative of?

- a. dominance
- b. codominance
- c. multiple alleles
- d. incomplete dominance

---

**Solution:**

D

**Exercise:**

**Problem:**

The ABO blood groups in humans are expressed as the  $I^A$ ,  $I^B$ , and  $i$  alleles. The  $I^A$  allele encodes the A blood group antigen,  $I^B$  encodes B, and  $i$  encodes O. Both A and B are dominant to O. If a heterozygous blood type A parent ( $I^A i$ ) and a heterozygous blood type B parent ( $I^B i$ ) mate, one quarter of their offspring will have AB blood type ( $I^A I^B$ ) in which both antigens are expressed equally. Therefore, ABO blood groups are an example of:

- a. multiple alleles and incomplete dominance
- b. codominance and incomplete dominance
- c. incomplete dominance only
- d. multiple alleles and codominance

---

**Solution:**

D

**Exercise:****Problem:**

In a mating between two individuals that are heterozygous for a recessive lethal allele that is expressed *in utero*, what genotypic ratio (homozygous dominant:heterozygous:homozygous recessive) would you expect to observe in the offspring?

- a. 1:2:1
- b. 3:1:1
- c. 1:2:0
- d. 0:2:1

---

**Solution:**

C

**Free Response****Exercise:****Problem:**

The gene for flower position in pea plants exists as axial or terminal alleles. Given that axial is dominant to terminal, list all of the possible  $F_1$  and  $F_2$  genotypes and phenotypes from a cross involving parents that are homozygous for each trait. Express genotypes with conventional genetic abbreviations.

---

**Solution:**

Because axial is dominant, the gene would be designated as  $A$ .  $F_1$  would be all heterozygous  $Aa$  with axial phenotype.  $F_2$  would have possible genotypes of  $AA$ ,  $Aa$ , and  $aa$ ; these would correspond to axial, axial, and terminal phenotypes, respectively.

**Exercise:**

**Problem:**

Use a Punnett square to predict the offspring in a cross between a dwarf pea plant (homozygous recessive) and a tall pea plant (heterozygous). What is the phenotypic ratio of the offspring?

---

**Solution:**

The Punnett square would be  $2 \times 2$  and will have  $T$  and  $T$  along the top, and  $T$  and  $t$  along the left side. Clockwise from the top left, the genotypes listed within the boxes will be  $Tt$ ,  $Tt$ ,  $tt$ , and  $tt$ . The phenotypic ratio will be 1 tall:1 dwarf.

**Exercise:**

**Problem:** Can a human male be a carrier of red-green color blindness?

---

**Solution:**

No, males can only express color blindness. They cannot carry it because an individual needs two X chromosomes to be a carrier.

**Glossary****allele**

gene variations that arise by mutation and exist at the same relative locations on homologous chromosomes

**autosomes**

any of the non-sex chromosomes

**codominance**

in a heterozygote, complete and simultaneous expression of both alleles for the same characteristic

**dominant lethal**

inheritance pattern in which an allele is lethal both in the homozygote and the heterozygote; this allele can only be transmitted if the lethality phenotype occurs after reproductive age

genotype

underlying genetic makeup, consisting of both physically visible and non-expressed alleles, of an organism

hemizygous

presence of only one allele for a characteristic, as in X-linkage; hemizyosity makes descriptions of dominance and recessiveness irrelevant

heterozygous

having two different alleles for a given gene on the homologous chromosome

homozygous

having two identical alleles for a given gene on the homologous chromosome

incomplete dominance

in a heterozygote, expression of two contrasting alleles such that the individual displays an intermediate phenotype

monohybrid

result of a cross between two true-breeding parents that express different traits for only one characteristic

phenotype

observable traits expressed by an organism

Punnett square

visual representation of a cross between two individuals in which the gametes of each individual are denoted along the top and side of a grid, respectively, and the possible zygotic genotypes are recombined at each box in the grid

recessive lethal

inheritance pattern in which an allele is only lethal in the homozygous form; the heterozygote may be normal or have some altered, non-lethal phenotype

sex-linked

any gene on a sex chromosome

test cross

cross between a dominant expressing individual with an unknown genotype and a homozygous recessive individual; the offspring phenotypes indicate whether the unknown parent is heterozygous or homozygous for the dominant trait

X-linked

gene present on the X, but not the Y chromosome

## Laws of Inheritance

By the end of this section, you will be able to:

- Explain Mendel's law of segregation and independent assortment in terms of genetics and the events of meiosis
- Use the forked-line method and the probability rules to calculate the probability of genotypes and phenotypes from multiple gene crosses
- Explain the effect of linkage and recombination on gamete genotypes
- Explain the phenotypic outcomes of epistatic effects between genes

Mendel generalized the results of his pea-plant experiments into four postulates, some of which are sometimes called “laws,” that describe the basis of dominant and recessive inheritance in diploid organisms. As you have learned, more complex extensions of Mendelism exist that do not exhibit the same  $F_2$  phenotypic ratios (3:1). Nevertheless, these laws summarize the basics of classical genetics.

## Pairs of Unit Factors, or Genes

Mendel proposed first that paired unit factors of heredity were transmitted faithfully from generation to generation by the dissociation and reassociation of paired factors during gametogenesis and fertilization, respectively. After he crossed peas with contrasting traits and found that the recessive trait resurfaced in the  $F_2$  generation, Mendel deduced that hereditary factors must be inherited as discrete units. This finding contradicted the belief at that time that parental traits were blended in the offspring.

## Alleles Can Be Dominant or Recessive

Mendel's **law of dominance** states that in a heterozygote, one trait will conceal the presence of another trait for the same characteristic. Rather than both alleles contributing to a phenotype, the dominant allele will be expressed exclusively. The recessive allele will remain “latent” but will be transmitted to offspring by the same manner in which the dominant allele is transmitted. The recessive trait will only be expressed by offspring that

have two copies of this allele ([link](#)), and these offspring will breed true when self-crossed.

Since Mendel's experiments with pea plants, other researchers have found that the law of dominance does not always hold true. Instead, several different patterns of inheritance have been found to exist.



The child in the photo expresses albinism, a recessive trait.

## Equal Segregation of Alleles

Observing that true-breeding pea plants with contrasting traits gave rise to  $F_1$  generations that all expressed the dominant trait and  $F_2$  generations that expressed the dominant and recessive traits in a 3:1 ratio, Mendel proposed the **law of segregation**. This law states that paired unit factors (genes) must segregate equally into gametes such that offspring have an equal likelihood



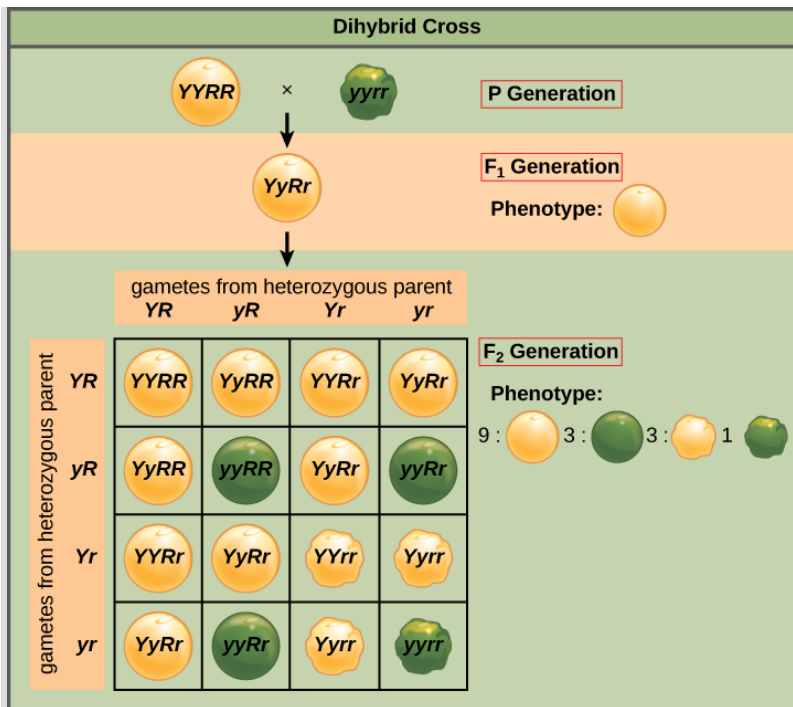
of inheriting either factor. For the  $F_2$  generation of a monohybrid cross, the following three possible combinations of genotypes could result: homozygous dominant, heterozygous, or homozygous recessive. Because heterozygotes could arise from two different pathways (receiving one dominant and one recessive allele from either parent), and because heterozygotes and homozygous dominant individuals are phenotypically identical, the law supports Mendel's observed 3:1 phenotypic ratio. The equal segregation of alleles is the reason we can apply the Punnett square to accurately predict the offspring of parents with known genotypes. The physical basis of Mendel's law of segregation is the first division of meiosis, in which the homologous chromosomes with their different versions of each gene are segregated into daughter nuclei. The role of the meiotic segregation of chromosomes in sexual reproduction was not understood by the scientific community during Mendel's lifetime.

## Independent Assortment

Mendel's **law of independent assortment** states that genes do not influence each other with regard to the sorting of alleles into gametes, and every possible combination of alleles for every gene is equally likely to occur. The independent assortment of genes can be illustrated by the **dihybrid** cross, a cross between two true-breeding parents that express different traits for two characteristics. Consider the characteristics of seed color and seed texture for two pea plants, one that has green, wrinkled seeds ( $yyrr$ ) and another that has yellow, round seeds ( $YYRR$ ). Because each parent is homozygous, the law of segregation indicates that the gametes for the green/wrinkled plant all are  $yr$ , and the gametes for the yellow/round plant are all  $YR$ . Therefore, the  $F_1$  generation of offspring all are  $YyRr$  ([link](#)).

### Note:

Art Connection



This dihybrid cross of pea plants involves the genes for seed color and texture.

In pea plants, purple flowers (P) are dominant to white flowers (p) and yellow peas (Y) are dominant to green peas (y). What are the possible genotypes and phenotypes for a cross between PpYY and ppYy pea plants? How many squares do you need to do a Punnett square analysis of this cross?

For the F<sub>2</sub> generation, the law of segregation requires that each gamete receive either an *R* allele or an *r* allele along with either a *Y* allele or a *y* allele. The law of independent assortment states that a gamete into which an *r* allele sorted would be equally likely to contain either a *Y* allele or a *y* allele. Thus, there are four equally likely gametes that can be formed when the *YyRr* heterozygote is self-crossed, as follows: *YR*, *Yr*, *yR*, and *yr*. Arranging these gametes along the top and left of a 4 × 4 Punnett square ([link](#)) gives us 16 equally likely genotypic combinations. From these genotypes, we infer a phenotypic ratio of 9 round/yellow:3 round/green:3

wrinkled/yellow:1 wrinkled/green ([link](#)). These are the offspring ratios we would expect, assuming we performed the crosses with a large enough sample size.

Because of independent assortment and dominance, the 9:3:3:1 dihybrid phenotypic ratio can be collapsed into two 3:1 ratios, characteristic of any monohybrid cross that follows a dominant and recessive pattern. Ignoring seed color and considering only seed texture in the above dihybrid cross, we would expect that three quarters of the  $F_2$  generation offspring would be round, and one quarter would be wrinkled. Similarly, isolating only seed color, we would assume that three quarters of the  $F_2$  offspring would be yellow and one quarter would be green. The sorting of alleles for texture and color are independent events, so we can apply the product rule.

Therefore, the proportion of round and yellow  $F_2$  offspring is expected to be  $(3/4) \times (3/4) = 9/16$ , and the proportion of wrinkled and green offspring is expected to be  $(1/4) \times (1/4) = 1/16$ . These proportions are identical to those obtained using a Punnett square. Round, green and wrinkled, yellow offspring can also be calculated using the product rule, as each of these genotypes includes one dominant and one recessive phenotype. Therefore, the proportion of each is calculated as  $(3/4) \times (1/4) = 3/16$ .

The law of independent assortment also indicates that a cross between yellow, wrinkled ( $YYrr$ ) and green, round ( $yyRR$ ) parents would yield the same  $F_1$  and  $F_2$  offspring as in the  $YYRR \times yyrr$  cross.

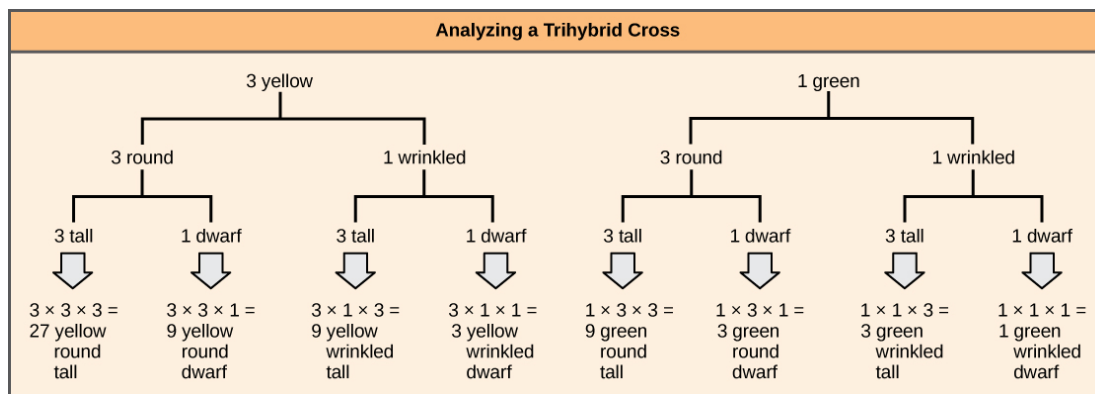
The physical basis for the law of independent assortment also lies in meiosis I, in which the different homologous pairs line up in random orientations. Each gamete can contain any combination of paternal and maternal chromosomes (and therefore the genes on them) because the orientation of tetrads on the metaphase plane is random.

## Forked-Line Method

When more than two genes are being considered, the Punnett-square method becomes unwieldy. For instance, examining a cross involving four genes would require a  $16 \times 16$  grid containing 256 boxes. It would be

extremely cumbersome to manually enter each genotype. For more complex crosses, the forked-line and probability methods are preferred.

To prepare a forked-line diagram for a cross between  $F_1$  heterozygotes resulting from a cross between  $AABBCC$  and  $aabbcc$  parents, we first create rows equal to the number of genes being considered, and then segregate the alleles in each row on forked lines according to the probabilities for individual monohybrid crosses ([link](#)). We then multiply the values along each forked path to obtain the  $F_2$  offspring probabilities. Note that this process is a diagrammatic version of the product rule. The values along each forked pathway can be multiplied because each gene assorts independently. For a trihybrid cross, the  $F_2$  phenotypic ratio is 27:9:9:9:3:3:3:1.



The forked-line method can be used to analyze a trihybrid cross. Here, the probability for color in the  $F_2$  generation occupies the top row (3 yellow:1 green). The probability for shape occupies the second row (3 round:1 wrinkled), and the probability for height occupies the third row (3 tall:1 dwarf).

The probability for each possible combination of traits is calculated by multiplying the probability for each individual trait. Thus, the probability of  $F_2$  offspring having yellow, round, and tall traits is  $3 \times 3 \times 3$ , or 27.

## Probability Method

While the forked-line method is a diagrammatic approach to keeping track of probabilities in a cross, the probability method gives the proportions of offspring expected to exhibit each phenotype (or genotype) without the added visual assistance. Both methods make use of the product rule and consider the alleles for each gene separately. Earlier, we examined the phenotypic proportions for a trihybrid cross using the forked-line method; now we will use the probability method to examine the genotypic proportions for a cross with even more genes.

For a trihybrid cross, writing out the forked-line method is tedious, albeit not as tedious as using the Punnett-square method. To fully demonstrate the power of the probability method, however, we can consider specific genetic calculations. For instance, for a tetrahybrid cross between individuals that are heterozygotes for all four genes, and in which all four genes are sorting independently and in a dominant and recessive pattern, what proportion of the offspring will be expected to be homozygous recessive for all four alleles? Rather than writing out every possible genotype, we can use the probability method. We know that for each gene, the fraction of homozygous recessive offspring will be  $1/4$ . Therefore, multiplying this fraction for each of the four genes,  $(1/4) \times (1/4) \times (1/4) \times (1/4)$ , we determine that  $1/256$  of the offspring will be quadruply homozygous recessive.

For the same tetrahybrid cross, what is the expected proportion of offspring that have the dominant phenotype at all four loci? We can answer this question using phenotypic proportions, but let's do it the hard way—using genotypic proportions. The question asks for the proportion of offspring that are 1) homozygous dominant at *A* or heterozygous at *A*, and 2) homozygous at *B* or heterozygous at *B*, and so on. Noting the “or” and “and” in each circumstance makes clear where to apply the sum and product rules. The probability of a homozygous dominant at *A* is  $1/4$  and the probability of a heterozygote at *A* is  $1/2$ . The probability of the homozygote or the heterozygote is  $1/4 + 1/2 = 3/4$  using the sum rule. The same probability can be obtained in the same way for each of the other genes, so that the probability of a dominant phenotype at *A* and *B* and *C* and

*D* is, using the product rule, equal to  $\frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4}$ , or 27/64. If you are ever unsure about how to combine probabilities, returning to the forked-line method should make it clear.

**Rules for Multihybrid Fertilization**

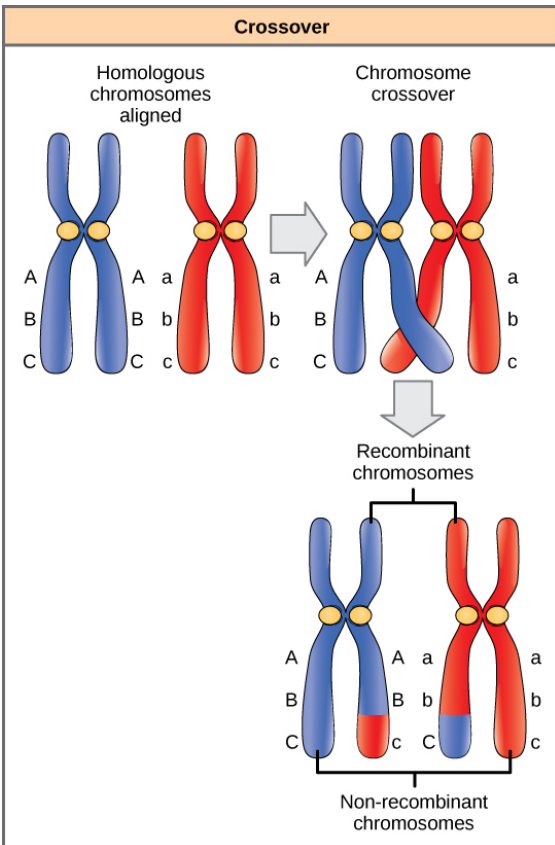
Predicting the genotypes and phenotypes of offspring from given crosses is the best way to test your knowledge of Mendelian genetics. Given a multihybrid cross that obeys independent assortment and follows a dominant and recessive pattern, several generalized rules exist; you can use these rules to check your results as you work through genetics calculations ([\[link\]](#)). To apply these rules, first you must determine *n*, the number of heterozygous gene pairs (the number of genes segregating two alleles each). For example, a cross between *AaBb* and *AaBb* heterozygotes has an *n* of 2. In contrast, a cross between *AABb* and *AABb* has an *n* of 1 because *A* is not heterozygous.

General Rules for Multihybrid Crosses	
General Rule	Number of Heterozygous Gene Pairs
Number of different F <sub>1</sub> gametes	2 <sup><i>n</i></sup>
Number of different F <sub>2</sub> genotypes	3 <sup><i>n</i></sup>
Given dominant and recessive inheritance, the number of different F <sub>2</sub> phenotypes	2 <sup><i>n</i></sup>

## Linked Genes Violate the Law of Independent Assortment

Although all of Mendel's pea characteristics behaved according to the law of independent assortment, we now know that some allele combinations are not inherited independently of each other. Genes that are located on separate non-homologous chromosomes will always sort independently. However, each chromosome contains hundreds or thousands of genes, organized linearly on chromosomes like beads on a string. The segregation of alleles into gametes can be influenced by **linkage**, in which genes that are located physically close to each other on the same chromosome are more likely to be inherited as a pair. However, because of the process of recombination, or "crossover," it is possible for two genes on the same chromosome to behave independently, or as if they are not linked. To understand this, let's consider the biological basis of gene linkage and recombination.

Homologous chromosomes possess the same genes in the same linear order. The alleles may differ on homologous chromosome pairs, but the genes to which they correspond do not. In preparation for the first division of meiosis, homologous chromosomes replicate and synapse. Like genes on the homologs align with each other. At this stage, segments of homologous chromosomes exchange linear segments of genetic material ([link](#)). This process is called recombination, or crossover, and it is a common genetic process. Because the genes are aligned during recombination, the gene order is not altered. Instead, the result of recombination is that maternal and paternal alleles are combined onto the same chromosome. Across a given chromosome, several recombination events may occur, causing extensive shuffling of alleles.



The process of crossover, or recombination, occurs when two homologous chromosomes align during meiosis and exchange a segment of genetic material. Here, the alleles for gene C were exchanged. The result is two recombinant and two non-recombinant chromosomes.

When two genes are located in close proximity on the same chromosome, they are considered linked, and their alleles tend to be transmitted through meiosis together. To exemplify this, imagine a dihybrid cross involving flower color and plant height in which the genes are next to each other on



the chromosome. If one homologous chromosome has alleles for tall plants and red flowers, and the other chromosome has genes for short plants and yellow flowers, then when the gametes are formed, the tall and red alleles will go together into a gamete and the short and yellow alleles will go into other gametes. These are called the parental genotypes because they have been inherited intact from the parents of the individual producing gametes. But unlike if the genes were on different chromosomes, there will be no gametes with tall and yellow alleles and no gametes with short and red alleles. If you create the Punnett square with these gametes, you will see that the classical Mendelian prediction of a 9:3:3:1 outcome of a dihybrid cross would not apply. As the distance between two genes increases, the probability of one or more crossovers between them increases, and the genes behave more like they are on separate chromosomes. Geneticists have used the proportion of recombinant gametes (the ones not like the parents) as a measure of how far apart genes are on a chromosome. Using this information, they have constructed elaborate maps of genes on chromosomes for well-studied organisms, including humans.

Mendel's seminal publication makes no mention of linkage, and many researchers have questioned whether he encountered linkage but chose not to publish those crosses out of concern that they would invalidate his independent assortment postulate. The garden pea has seven chromosomes, and some have suggested that his choice of seven characteristics was not a coincidence. However, even if the genes he examined were not located on separate chromosomes, it is possible that he simply did not observe linkage because of the extensive shuffling effects of recombination.

**Note:**

Scientific Method Connection

**Testing the Hypothesis of Independent Assortment**

To better appreciate the amount of labor and ingenuity that went into Mendel's experiments, proceed through one of Mendel's dihybrid crosses.

**Question:** What will be the offspring of a dihybrid cross?

**Background:** Consider that pea plants mature in one growing season, and you have access to a large garden in which you can cultivate thousands of

pea plants. There are several true-breeding plants with the following pairs of traits: tall plants with inflated pods, and dwarf plants with constricted pods. Before the plants have matured, you remove the pollen-producing organs from the tall/inflated plants in your crosses to prevent self-fertilization. Upon plant maturation, the plants are manually crossed by transferring pollen from the dwarf/constricted plants to the stigmata of the tall/inflated plants.

**Hypothesis:** Both trait pairs will sort independently according to Mendelian laws. When the true-breeding parents are crossed, all of the  $F_1$  offspring are tall and have inflated pods, which indicates that the tall and inflated traits are dominant over the dwarf and constricted traits, respectively. A self-cross of the  $F_1$  heterozygotes results in 2,000  $F_2$  progeny.

**Test the hypothesis:** Because each trait pair sorts independently, the ratios of tall:dwarf and inflated:constricted are each expected to be 3:1. The tall/dwarf trait pair is called  $T/t$ , and the inflated/constricted trait pair is designated  $I/i$ . Each member of the  $F_1$  generation therefore has a genotype of  $TtIi$ . Construct a grid analogous to [\[link\]](#), in which you cross two  $TtIi$  individuals. Each individual can donate four combinations of two traits:  $TI$ ,  $Ti$ ,  $tI$ , or  $ti$ , meaning that there are 16 possibilities of offspring genotypes. Because the  $T$  and  $I$  alleles are dominant, any individual having one or two of those alleles will express the tall or inflated phenotypes, respectively, regardless if they also have a  $t$  or  $i$  allele. Only individuals that are  $tt$  or  $ii$  will express the dwarf and constricted alleles, respectively. As shown in [\[link\]](#), you predict that you will observe the following offspring proportions: tall/inflated:tall/constricted:dwarf/inflated:dwarf/constricted in a 9:3:3:1 ratio. Notice from the grid that when considering the tall/dwarf and inflated/constricted trait pairs in isolation, they are each inherited in 3:1 ratios.

		<i>TtIi</i>			
		<i>TI</i>	<i>Ti</i>	<i>tI</i>	<i>ti</i>
<i>TtIi</i>	<i>TI</i>	<i>TTII</i>	<i>TTIi</i>	<i>TtII</i>	<i>TtIi</i>
	<i>Ti</i>	<i>TTIi</i>	<i>TTii</i>	<i>TtIi</i>	<i>Ttii</i>
	<i>tI</i>	<i>TtII</i>	<i>TtIi</i>	<i>ttII</i>	<i>ttIi</i>
	<i>ti</i>	<i>TtIi</i>	<i>Ttii</i>	<i>ttIi</i>	<i>ttii</i>

This figure shows all possible combinations of offspring resulting from a dihybrid cross of pea plants that are heterozygous for the tall/dwarf and inflated/constricted alleles.

**Test the hypothesis:** You cross the dwarf and tall plants and then self-cross the offspring. For best results, this is repeated with hundreds or even thousands of pea plants. What special precautions should be taken in the crosses and in growing the plants?

**Analyze your data:** You observe the following plant phenotypes in the F<sub>2</sub> generation: 2706 tall/inflated, 930 tall/constricted, 888 dwarf/inflated, and 300 dwarf/constricted. Reduce these findings to a ratio and determine if they are consistent with Mendelian laws.

**Form a conclusion:** Were the results close to the expected 9:3:3:1 phenotypic ratio? Do the results support the prediction? What might be observed if far fewer plants were used, given that alleles segregate randomly into gametes? Try to imagine growing that many pea plants, and consider the potential for experimental error. For instance, what would happen if it was extremely windy one day?

## Epistasis

Mendel's studies in pea plants implied that the sum of an individual's phenotype was controlled by genes (or as he called them, unit factors), such that every characteristic was distinctly and completely controlled by a single gene. In fact, single observable characteristics are almost always under the influence of multiple genes (each with two or more alleles) acting in unison. For example, at least eight genes contribute to eye color in humans.

### Note:

#### Link to Learning



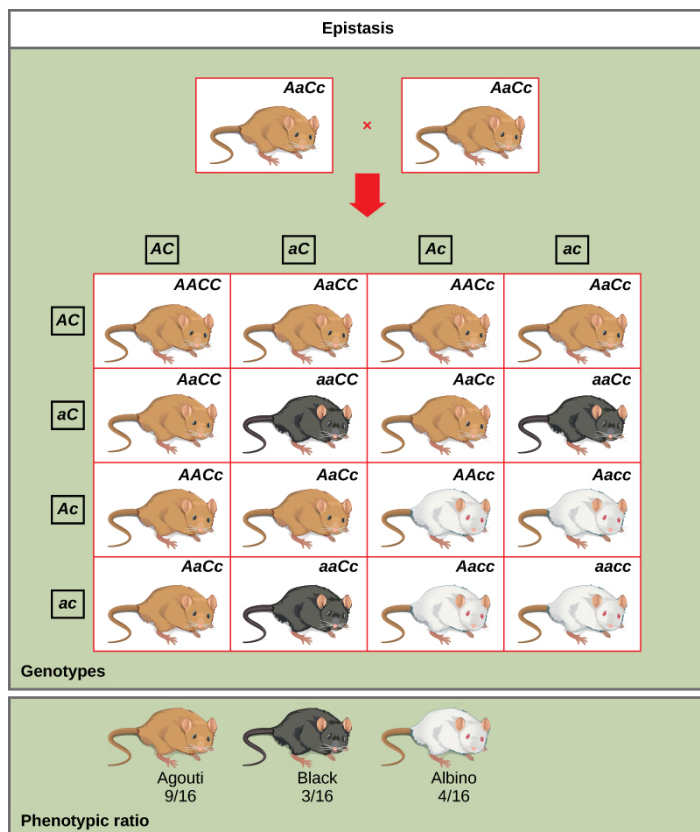
Eye color in humans is determined by multiple genes. Use the [Eye Color Calculator](#) to predict the eye color of children from parental eye color.

In some cases, several genes can contribute to aspects of a common phenotype without their gene products ever directly interacting. In the case of organ development, for instance, genes may be expressed sequentially, with each gene adding to the complexity and specificity of the organ. Genes may function in complementary or synergistic fashions, such that two or more genes need to be expressed simultaneously to affect a phenotype. Genes may also oppose each other, with one gene modifying the expression of another.

In **epistasis**, the interaction between genes is antagonistic, such that one gene masks or interferes with the expression of another. “Epistasis” is a

word composed of Greek roots that mean “standing upon.” The alleles that are being masked or silenced are said to be hypostatic to the epistatic alleles that are doing the masking. Often the biochemical basis of epistasis is a gene pathway in which the expression of one gene is dependent on the function of a gene that precedes or follows it in the pathway.

An example of epistasis is pigmentation in mice. The wild-type coat color, agouti (*AA*), is dominant to solid-colored fur (*aa*). However, a separate gene (*C*) is necessary for pigment production. A mouse with a recessive *c* allele at this locus is unable to produce pigment and is albino regardless of the allele present at locus *A* ([link](#)). Therefore, the genotypes *AAcc*, *Aacc*, and *aacc* all produce the same albino phenotype. A cross between heterozygotes for both genes (*AaCc* x *AaCc*) would generate offspring with a phenotypic ratio of 9 agouti:3 solid color:4 albino ([link](#)). In this case, the *C* gene is epistatic to the *A* gene.



In mice, the mottled agouti coat color (A) is dominant to a solid coloration, such as black or gray. A gene at a separate locus (C) is responsible for pigment production. The recessive *c* allele does not produce pigment, and a mouse with the homozygous recessive *cc* genotype is albino regardless of the allele present at the A locus. Thus, the C gene is epistatic to the A gene.

Epistasis can also occur when a dominant allele masks expression at a separate gene. Fruit color in summer squash is expressed in this way. Homozygous recessive expression of the *W* gene (*ww*) coupled with homozygous dominant or heterozygous expression of the *Y* gene (*YY* or *Yy*) generates yellow fruit, and the *wwyy* genotype produces green fruit. However, if a dominant copy of the *W* gene is present in the homozygous or heterozygous form, the summer squash will produce white fruit regardless of the *Y* alleles. A cross between white heterozygotes for both genes (*WwYy* × *WwYy*) would produce offspring with a phenotypic ratio of 12 white:3 yellow:1 green.

Finally, epistasis can be reciprocal such that either gene, when present in the dominant (or recessive) form, expresses the same phenotype. In the shepherd's purse plant (*Capsella bursa-pastoris*), the characteristic of seed shape is controlled by two genes in a dominant epistatic relationship. When the genes *A* and *B* are both homozygous recessive (*aabb*), the seeds are ovoid. If the dominant allele for either of these genes is present, the result is triangular seeds. That is, every possible genotype other than *aabb* results in triangular seeds, and a cross between heterozygotes for both genes (*AaBb* × *AaBb*) would yield offspring with a phenotypic ratio of 15 triangular:1 ovoid.

As you work through genetics problems, keep in mind that any single characteristic that results in a phenotypic ratio that totals 16 is typical of a

two-gene interaction. Recall the phenotypic inheritance pattern for Mendel's dihybrid cross, which considered two non-interacting genes—9:3:3:1. Similarly, we would expect interacting gene pairs to also exhibit ratios expressed as 16 parts. Note that we are assuming the interacting genes are not linked; they are still assorting independently into gametes.

**Note:**

Link to Learning



For an excellent review of Mendel's experiments and to perform your own crosses and identify patterns of inheritance, visit the [Mendel's Peas](#) web lab.

## Section Summary

Mendel postulated that genes (characteristics) are inherited as pairs of alleles (traits) that behave in a dominant and recessive pattern. Alleles segregate into gametes such that each gamete is equally likely to receive either one of the two alleles present in a diploid individual. In addition, genes are assorted into gametes independently of one another. That is, alleles are generally not more likely to segregate into a gamete with a particular allele of another gene. A dihybrid cross demonstrates independent assortment when the genes in question are on different chromosomes or distant from each other on the same chromosome. For crosses involving more than two genes, use the forked line or probability methods to predict offspring genotypes and phenotypes rather than a Punnett square.

Although chromosomes sort independently into gametes during meiosis, Mendel's law of independent assortment refers to genes, not chromosomes, and a single chromosome may carry more than 1,000 genes. When genes are located in close proximity on the same chromosome, their alleles tend to be inherited together. This results in offspring ratios that violate Mendel's law of independent assortment. However, recombination serves to exchange genetic material on homologous chromosomes such that maternal and paternal alleles may be recombined on the same chromosome. This is why alleles on a given chromosome are not always inherited together. Recombination is a random event occurring anywhere on a chromosome. Therefore, genes that are far apart on the same chromosome are likely to still assort independently because of recombination events that occurred in the intervening chromosomal space.

Whether or not they are sorting independently, genes may interact at the level of gene products such that the expression of an allele for one gene masks or modifies the expression of an allele for a different gene. This is called epistasis.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) In pea plants, purple flowers (P) are dominant to white flowers (p) and yellow peas (Y) are dominant to green peas (y). What are the possible genotypes and phenotypes for a cross between PpYY and ppYy pea plants? How many squares do you need to do a Punnett square analysis of this cross?

---

#### Solution:

[\[link\]](#) The possible genotypes are PpYY, PpYy, ppYY, and ppYy. The former two genotypes would result in plants with purple flowers and yellow peas, while the latter two genotypes would result in plants with white flowers with yellow peas, for a 1:1 ratio of each phenotype. You



only need a  $2 \times 2$  Punnett square (four squares total) to do this analysis because two of the alleles are homozygous.

## Multiple Choice

### Exercise:

#### Problem:

Assuming no gene linkage, in a dihybrid cross of  $AABB \times aabb$  with  $AaBb$   $F_1$  heterozygotes, what is the ratio of the  $F_1$  gametes ( $AB$ ,  $aB$ ,  $Ab$ ,  $ab$ ) that will give rise to the  $F_2$  offspring?

- a. 1:1:1:1
- b. 1:3:3:1
- c. 1:2:2:1
- d. 4:3:2:1

---

#### Solution:

A

### Exercise:

#### Problem:

The forked line and probability methods make use of what probability rule?

- a. test cross
- b. product rule
- c. monohybrid rule
- d. sum rule

---

#### Solution:

B

**Exercise:****Problem:**

How many different offspring genotypes are expected in a trihybrid cross between parents heterozygous for all three traits when the traits behave in a dominant and recessive pattern? How many phenotypes?

- a. 64 genotypes; 16 phenotypes
- b. 16 genotypes; 64 phenotypes
- c. 8 genotypes; 27 phenotypes
- d. 27 genotypes; 8 phenotypes

---

**Solution:**

D

**Free Response****Exercise:****Problem:**

Use the probability method to calculate the genotypes and genotypic proportions of a cross between *AABBCc* and *Aabbcc* parents.

---

**Solution:**

Considering each gene separately, the cross at *A* will produce offspring of which half are *AA* and half are *Aa*; *B* will produce all *Bb*; *C* will produce half *Cc* and half *cc*. Proportions then are  $(1/2) \times (1) \times (1/2)$ , or  $1/4$  *AABbCc*; continuing for the other possibilities yields  $1/4$  *AABbcc*,  $1/4$  *AaBbCc*, and  $1/4$  *AaBbcc*. The proportions therefore are 1:1:1:1.

**Exercise:**

**Problem:**

Explain epistasis in terms of its Greek-language roots “standing upon.”

---

**Solution:**

Epistasis describes an antagonistic interaction between genes wherein one gene masks or interferes with the expression of another. The gene that is interfering is referred to as epistatic, as if it is “standing upon” the other (hypostatic) gene to block its expression.

**Exercise:****Problem:**

In Section 12.3, “Laws of Inheritance,” an example of epistasis was given for the summer squash. Cross white  $WwYy$  heterozygotes to prove the phenotypic ratio of 12 white:3 yellow:1 green that was given in the text.

---

**Solution:**

The cross can be represented as a  $4 \times 4$  Punnett square, with the following gametes for each parent:  $WY$ ,  $Wy$ ,  $wY$ , and  $wy$ . For all 12 of the offspring that express a dominant  $W$  gene, the offspring will be white. The three offspring that are homozygous recessive for  $w$  but express a dominant  $Y$  gene will be yellow. The remaining  $wwyy$  offspring will be green.

**Glossary****dihybrid**

result of a cross between two true-breeding parents that express different traits for two characteristics

**epistasis**

antagonistic interaction between genes such that one gene masks or interferes with the expression of another

law of dominance

in a heterozygote, one trait will conceal the presence of another trait for the same characteristic

law of independent assortment

genes do not influence each other with regard to sorting of alleles into gametes; every possible combination of alleles is equally likely to occur

law of segregation

paired unit factors (i.e., genes) segregate equally into gametes such that offspring have an equal likelihood of inheriting any combination of factors

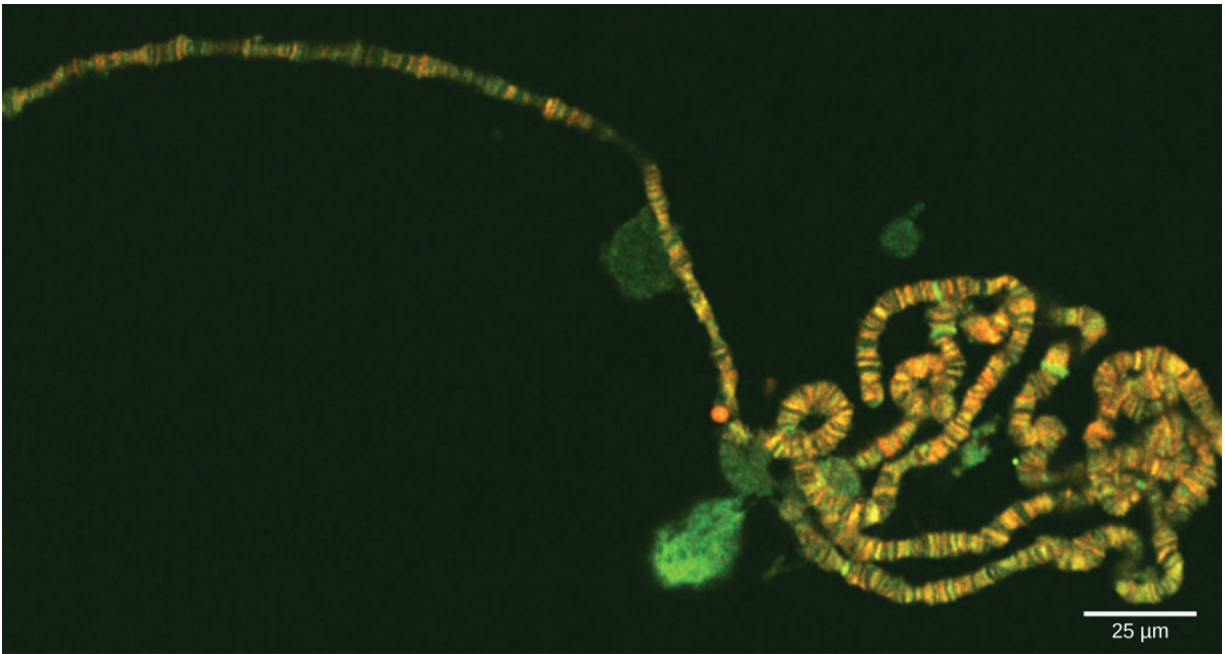
linkage

phenomenon in which alleles that are located in close proximity to each other on the same chromosome are more likely to be inherited together

## Introduction

class="introduction"

Chromosomes are threadlike nuclear structures consisting of DNA and proteins that serve as the repositories for genetic information. The chromosomes depicted here were isolated from a fruit fly's salivary gland, stained with dye, and visualized under a microscope. Akin to miniature bar codes, chromosomes absorb different dyes to produce characteristic banding patterns, which allows for their routine identification. (credit: modification of work by "LPLT"/Wikimedia Commons; scale-bar data from Matt Russell)



The gene is the physical unit of inheritance, and genes are arranged in a linear order on chromosomes. The behaviors and interactions of chromosomes during meiosis explain, at a cellular level, the patterns of inheritance that we observe in populations. Genetic disorders involving alterations in chromosome number or structure may have dramatic effects and can prevent a fertilized egg from developing altogether.

## Chromosomal Theory and Genetic Linkage

By the end of this section, you will be able to:

- Discuss Sutton's Chromosomal Theory of Inheritance
- Describe genetic linkage
- Explain the process of homologous recombination, or crossing over
- Describe how chromosome maps are created
- Calculate the distances between three genes on a chromosome using a three-point test cross

Long before chromosomes were visualized under a microscope, the father of modern genetics, Gregor Mendel, began studying heredity in 1843. With the improvement of microscopic techniques during the late 1800s, cell biologists could stain and visualize subcellular structures with dyes and observe their actions during cell division and meiosis. With each mitotic division, chromosomes replicated, condensed from an amorphous (no constant shape) nuclear mass into distinct X-shaped bodies (pairs of identical sister chromatids), and migrated to separate cellular poles.

## Chromosomal Theory of Inheritance

The speculation that chromosomes might be the key to understanding heredity led several scientists to examine Mendel's publications and re-evaluate his model in terms of the behavior of chromosomes during mitosis and meiosis. In 1902, Theodor Boveri observed that proper embryonic development of sea urchins does not occur unless chromosomes are present. That same year, Walter Sutton observed the separation of chromosomes into daughter cells during meiosis ([\[link\]](#)). Together, these observations led to the development of the **Chromosomal Theory of Inheritance**, which identified chromosomes as the genetic material responsible for Mendelian inheritance.



(a) Walter Sutton and (b) Theodor Boveri are credited with developing the Chromosomal Theory of Inheritance, which states that chromosomes carry the unit of heredity (genes).

The Chromosomal Theory of Inheritance was consistent with Mendel's laws and was supported by the following observations:

- During meiosis, homologous chromosome pairs migrate as discrete structures that are independent of other chromosome pairs.
- The sorting of chromosomes from each homologous pair into pre-gametes appears to be random.
- Each parent synthesizes gametes that contain only half of their chromosomal complement.
- Even though male and female gametes (sperm and egg) differ in size and morphology, they have the same number of chromosomes, suggesting equal genetic contributions from each parent.
- The gametic chromosomes combine during fertilization to produce offspring with the same chromosome number as their parents.

Despite compelling correlations between the behavior of chromosomes during meiosis and Mendel's abstract laws, the Chromosomal Theory of



Inheritance was proposed long before there was any direct evidence that traits were carried on chromosomes. Critics pointed out that individuals had far more independently segregating traits than they had chromosomes. It was only after several years of carrying out crosses with the fruit fly, *Drosophila melanogaster*, that Thomas Hunt Morgan provided experimental evidence to support the Chromosomal Theory of Inheritance.

## Genetic Linkage and Distances

Mendel's work suggested that traits are inherited independently of each other. Morgan identified a 1:1 correspondence between a segregating trait and the X chromosome, suggesting that the random segregation of chromosomes was the physical basis of Mendel's model. This also demonstrated that linked genes disrupt Mendel's predicted outcomes. The fact that each chromosome can carry many linked genes explains how individuals can have many more traits than they have chromosomes. However, observations by researchers in Morgan's laboratory suggested that alleles positioned on the same chromosome were not always inherited together. During meiosis, linked genes somehow became unlinked.

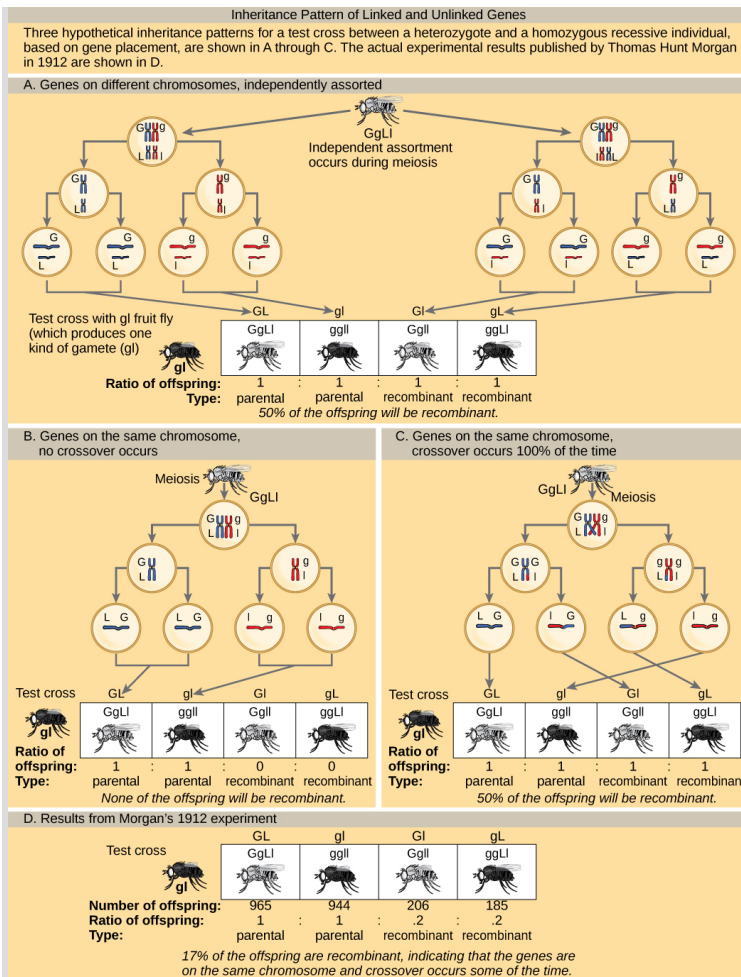
## Homologous Recombination

In 1909, Frans Janssen observed chiasmata—the point at which chromatids are in contact with each other and may exchange segments—prior to the first division of meiosis. He suggested that alleles become unlinked and chromosomes physically exchange segments. As chromosomes condensed and paired with their homologs, they appeared to interact at distinct points. Janssen suggested that these points corresponded to regions in which chromosome segments were exchanged. It is now known that the pairing and interaction between homologous chromosomes, known as synapsis, does more than simply organize the homologs for migration to separate daughter cells. When synapsed, homologous chromosomes undergo reciprocal physical exchanges at their arms in a process called **homologous recombination**, or more simply, “crossing over.”

To better understand the type of experimental results that researchers were obtaining at this time, consider a heterozygous individual that inherited dominant maternal alleles for two genes on the same chromosome (such as *AB*) and two recessive paternal alleles for those same genes (such as *ab*). If the genes are linked, one would expect this individual to produce gametes that are either *AB* or *ab* with a 1:1 ratio. If the genes are unlinked, the individual should produce *AB*, *Ab*, *aB*, and *ab* gametes with equal frequencies, according to the Mendelian concept of independent assortment. Because they correspond to new allele combinations, the genotypes *Ab* and *aB* are **nonparental types** that result from homologous recombination during meiosis. **Parental types** are progeny that exhibit the same allelic combination as their parents. Morgan and his colleagues, however, found that when such heterozygous individuals were test crossed to a homozygous recessive parent (*AaBb* × *aabb*), both parental and nonparental cases occurred. For example, 950 offspring might be recovered that were either *AaBb* or *aabb*, but 50 offspring would also be obtained that were either *Aabb* or *aaBb*. These results suggested that linkage occurred most often, but a significant minority of offspring were the products of recombination.

**Note:**

Art Connection



Inheritance patterns of unlinked and linked genes are shown. In (a), two genes are located on different chromosomes so independent assortment occurs during meiosis. The offspring have an equal chance of being the parental type (inheriting the same combination of traits as the parents) or a nonparental type (inheriting a different combination of traits than the parents). In (b), two genes are very close together on the same chromosome so that no crossing over occurs between them. The genes are therefore always inherited together and all of the offspring are the

parental type. In (c), two genes are far apart on the chromosome such that crossing over occurs during every meiotic event. The recombination frequency will be the same as if the genes were on separate chromosomes. (d) The actual recombination frequency of fruit fly wing length and body color that Thomas Morgan observed in 1912 was 17 percent. A crossover frequency between 0 percent and 50 percent indicates that the genes are on the same chromosome and crossover occurs some of the time.

In a test cross for two characteristics such as the one shown here, can the predicted frequency of recombinant offspring be 60 percent? Why or why not?

## Genetic Maps

Janssen did not have the technology to demonstrate crossing over so it remained an abstract idea that was not widely accepted. Scientists thought chiasmata were a variation on synapsis and could not understand how chromosomes could break and rejoin. Yet, the data were clear that linkage did not always occur. Ultimately, it took a young undergraduate student and an “all-nighter” to mathematically elucidate the problem of linkage and recombination.

In 1913, Alfred Sturtevant, a student in Morgan’s laboratory, gathered results from researchers in the laboratory, and took them home one night to mull them over. By the next morning, he had created the first “chromosome map,” a linear representation of gene order and relative distance on a chromosome ([link](#)).

## Note:

### Art Connection

Genetic Map Based on Recombination Frequencies in <i>Drosophila</i>			
MUTANT			WILD TYPE
Short aristae	0		Long aristae
Black body	48.5		Gray body
Cinnabar eyes	57.5		Red eyes
Vestigial wings	65.5		Normal wings
Brown eyes	104.5		Red eyes
Values in centimorgan (cM) map units; recombination frequency of 0.01 = 1 cM			

This genetic map orders *Drosophila* genes on the basis of recombination frequency.

Which of the following statements is true?

- a. Recombination of the body color and red/cinnabar eye alleles will occur more frequently than recombination of the alleles for wing length and aristae length.
- b. Recombination of the body color and aristae length alleles will occur more frequently than recombination of red/brown eye alleles and the aristae length alleles.
- c. Recombination of the gray/black body color and long/short aristae alleles will not occur.
- d. Recombination of the red/brown eye and long/short aristae alleles will occur more frequently than recombination of the alleles for wing length and body color.

As shown in [\[link\]](#), by using recombination frequency to predict genetic distance, the relative order of genes on chromosome 2 could be inferred. The values shown represent map distances in centimorgans (cM), which correspond to recombination frequencies (in percent). Therefore, the genes for body color and wing size were  $65.5 - 48.5 = 17$  cM apart, indicating that the maternal and paternal alleles for these genes recombine in 17 percent of offspring, on average.

To construct a chromosome map, Sturtevant assumed that genes were ordered serially on threadlike chromosomes. He also assumed that the incidence of recombination between two homologous chromosomes could occur with equal likelihood anywhere along the length of the chromosome. Operating under these assumptions, Sturtevant postulated that alleles that were far apart on a chromosome were more likely to dissociate during meiosis simply because there was a larger region over which recombination could occur. Conversely, alleles that were close to each other on the chromosome were likely to be inherited together. The average number of crossovers between two alleles—that is, their **recombination frequency**—correlated with their genetic distance from each other, relative to the locations of other genes on that chromosome. Considering the example cross between *AaBb* and *aabb* above, the frequency of recombination could be calculated as  $50/1000 = 0.05$ . That is, the likelihood of a crossover between genes *A/a* and *B/b* was 0.05, or 5 percent. Such a result would indicate that the genes were definitively linked, but that they were far enough apart for crossovers to occasionally occur. Sturtevant divided his genetic map into map units, or **centimorgans (cM)**, in which a recombination frequency of 0.01 corresponds to 1 cM.

By representing alleles in a linear map, Sturtevant suggested that genes can range from being perfectly linked (recombination frequency = 0) to being perfectly unlinked (recombination frequency = 0.5) when genes are on different chromosomes or genes are separated very far apart on the same chromosome. Perfectly unlinked genes correspond to the frequencies predicted by Mendel to assort independently in a dihybrid cross. A recombination frequency of 0.5 indicates that 50 percent of offspring are recombinants and the other 50 percent are parental types. That is, every type of allele combination is represented with equal frequency. This

representation allowed Sturtevant to additively calculate distances between several genes on the same chromosome. However, as the genetic distances approached 0.50, his predictions became less accurate because it was not clear whether the genes were very far apart on the same chromosome or on different chromosomes.

In 1931, Barbara McClintock and Harriet Creighton demonstrated the crossover of homologous chromosomes in corn plants. Weeks later, homologous recombination in *Drosophila* was demonstrated microscopically by Curt Stern. Stern observed several X-linked phenotypes that were associated with a structurally unusual and dissimilar X chromosome pair in which one X was missing a small terminal segment, and the other X was fused to a piece of the Y chromosome. By crossing flies, observing their offspring, and then visualizing the offspring's chromosomes, Stern demonstrated that every time the offspring allele combination deviated from either of the parental combinations, there was a corresponding exchange of an X chromosome segment. Using mutant flies with structurally distinct X chromosomes was the key to observing the products of recombination because DNA sequencing and other molecular tools were not yet available. It is now known that homologous chromosomes regularly exchange segments in meiosis by reciprocally breaking and rejoining their DNA at precise locations.

**Note:**

Link to Learning



Review Sturtevant's process to create a genetic map on the basis of recombination frequencies [here](#).

## **Mendel's Mapped Traits**

Homologous recombination is a common genetic process, yet Mendel never observed it. Had he investigated both linked and unlinked genes, it would have been much more difficult for him to create a unified model of his data on the basis of probabilistic calculations. Researchers who have since mapped the seven traits investigated by Mendel onto the seven chromosomes of the pea plant genome have confirmed that all of the genes he examined are either on separate chromosomes or are sufficiently far apart as to be statistically unlinked. Some have suggested that Mendel was enormously lucky to select only unlinked genes, whereas others question whether Mendel discarded any data suggesting linkage. In any case, Mendel consistently observed independent assortment because he examined genes that were effectively unlinked.

## **Section Summary**

The Chromosomal Theory of inheritance, proposed by Sutton and Boveri, states that chromosomes are the vehicles of genetic heredity. Neither Mendelian genetics nor gene linkage is perfectly accurate; instead, chromosome behavior involves segregation, independent assortment, and occasionally, linkage. Sturtevant devised a method to assess recombination frequency and infer the relative positions and distances of linked genes on a chromosome on the basis of the average number of crossovers in the intervening region between the genes. Sturtevant correctly presumed that genes are arranged in serial order on chromosomes and that recombination between homologs can occur anywhere on a chromosome with equal likelihood. Whereas linkage causes alleles on the same chromosome to be inherited together, homologous recombination biases alleles toward an inheritance pattern of independent assortment.

## **Art Connections**

### **Exercise:**



**Problem:**

[\[link\]](#) In a test cross for two characteristics such as the one shown here, can the predicted frequency of recombinant offspring be 60 percent? Why or why not?

---

**Solution:**

[\[link\]](#) No. The predicted frequency of recombinant offspring ranges from 0% (for linked traits) to 50% (for unlinked traits).

**Exercise:**

**Problem:** [\[link\]](#) Which of the following statements is true?

- a. Recombination of the body color and red/cinnabar eye alleles will occur more frequently than recombination of the alleles for wing length and aristae length.
  - b. Recombination of the body color and aristae length alleles will occur more frequently than recombination of red/brown eye alleles and the aristae length alleles.
  - c. Recombination of the gray/black body color and long/short aristae alleles will not occur.
  - d. Recombination of the red/brown eye and long/short aristae alleles will occur more frequently than recombination of the alleles for wing length and body color.
- 

**Solution:**

[\[link\]](#) D

**Review Questions****Exercise:**

**Problem:**

X-linked recessive traits in humans (or in *Drosophila*) are observed \_\_\_\_\_.

- a. in more males than females
- b. in more females than males
- c. in males and females equally
- d. in different distributions depending on the trait

---

**Solution:**

A

**Exercise:****Problem:**

The first suggestion that chromosomes may physically exchange segments came from the microscopic identification of \_\_\_\_\_.

- a. synapsis
- b. sister chromatids
- c. chiasmata
- d. alleles

---

**Solution:**

C

**Exercise:****Problem:**

Which recombination frequency corresponds to independent assortment and the absence of linkage?

- a. 0

- b. 0.25
- c. 0.50
- d. 0.75

---

**Solution:**

C

**Exercise:**

**Problem:**

Which recombination frequency corresponds to perfect linkage and violates the law of independent assortment?

- a. 0
- b. 0.25
- c. 0.50
- d. 0.75

---

**Solution:**

A

**Free Response**

**Exercise:**

**Problem:**

Explain how the Chromosomal Theory of Inheritance helped to advance our understanding of genetics.

---

**Solution:**

The Chromosomal Theory of Inheritance proposed that genes reside on chromosomes. The understanding that chromosomes are linear arrays

of genes explained linkage, and crossing over explained recombination.

## **Glossary**

centimorgan (cM)

(also, map unit) relative distance that corresponds to a recombination frequency of 0.01

Chromosomal Theory of Inheritance

theory proposing that chromosomes are the vehicles of genes and that their behavior during meiosis is the physical basis of the inheritance patterns that Mendel observed

homologous recombination

process by which homologous chromosomes undergo reciprocal physical exchanges at their arms, also known as crossing over

nonparental (recombinant) type

progeny resulting from homologous recombination that exhibits a different allele combination compared with its parents

parental types

progeny that exhibits the same allelic combination as its parents

recombination frequency

average number of crossovers between two alleles; observed as the number of nonparental types in a population of progeny

## Chromosomal Basis of Inherited Disorders

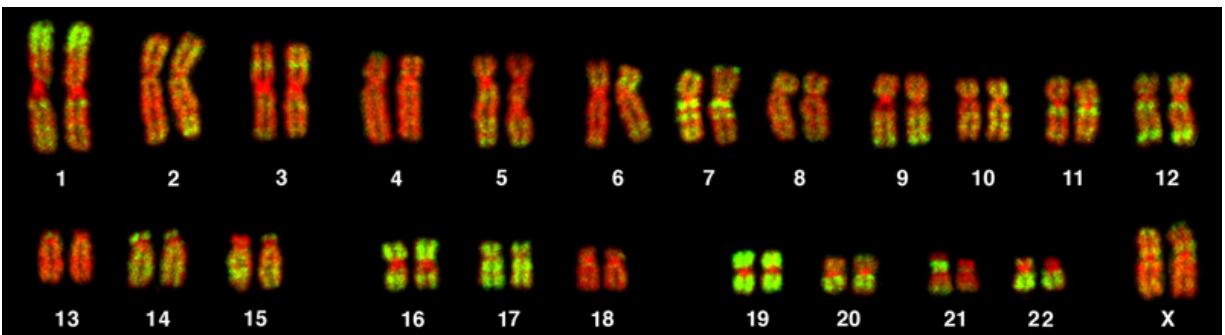
By the end of this section, you will be able to:

- Describe how a karyogram is created
- Explain how nondisjunction leads to disorders in chromosome number
- Compare disorders caused by aneuploidy
- Describe how errors in chromosome structure occur through inversions and translocations

Inherited disorders can arise when chromosomes behave abnormally during meiosis. Chromosome disorders can be divided into two categories: abnormalities in chromosome number and chromosomal structural rearrangements. Because even small segments of chromosomes can span many genes, chromosomal disorders are characteristically dramatic and often fatal.

### Identification of Chromosomes

The isolation and microscopic observation of chromosomes forms the basis of cytogenetics and is the primary method by which clinicians detect chromosomal abnormalities in humans. A **karyotype** is the number and appearance of chromosomes, and includes their length, banding pattern, and centromere position. To obtain a view of an individual's karyotype, cytologists photograph the chromosomes and then cut and paste each chromosome into a chart, or **karyogram**, also known as an ideogram ([\[link\]](#)).



This karyotype is of a female human. Notice that homologous chromosomes are the same size, and have the same centromere positions and banding patterns. A human male would have an XY chromosome pair instead of the XX pair shown. (credit: Andreas Blozer et al)

In a given species, chromosomes can be identified by their number, size, centromere position, and banding pattern. In a human karyotype, **autosomes** or “body chromosomes” (all of the non–sex chromosomes) are generally organized in approximate order of size from largest (chromosome 1) to smallest (chromosome 22). The X and Y chromosomes are not autosomes. However, chromosome 21 is actually shorter than chromosome 22. This was discovered after the naming of Down syndrome as trisomy 21, reflecting how this disease results from possessing one extra chromosome 21 (three total). Not wanting to change the name of this important disease, chromosome 21 retained its numbering, despite describing the shortest set of chromosomes. The chromosome “arms” projecting from either end of the centromere may be designated as short or long, depending on their relative lengths. The short arm is abbreviated *p* (for “petite”), whereas the long arm is abbreviated *q* (because it follows “p” alphabetically). Each arm is further subdivided and denoted by a number. Using this naming system, locations on chromosomes can be described consistently in the scientific literature.

**Note:****Career Connection****Geneticists Use Karyograms to Identify Chromosomal Aberrations**

Although Mendel is referred to as the “father of modern genetics,” he performed his experiments with none of the tools that the geneticists of today routinely employ. One such powerful cytological technique is karyotyping, a method in which traits characterized by chromosomal abnormalities can be identified from a single cell. To observe an individual’s karyotype, a person’s cells (like white blood cells) are first collected from a blood sample or other tissue. In the laboratory, the isolated

cells are stimulated to begin actively dividing. A chemical called colchicine is then applied to cells to arrest condensed chromosomes in metaphase. Cells are then made to swell using a hypotonic solution so the chromosomes spread apart. Finally, the sample is preserved in a fixative and applied to a slide.

The geneticist then stains chromosomes with one of several dyes to better visualize the distinct and reproducible banding patterns of each chromosome pair. Following staining, the chromosomes are viewed using bright-field microscopy. A common stain choice is the Giemsa stain.

Giemsa staining results in approximately 400–800 bands (of tightly coiled DNA and condensed proteins) arranged along all of the 23 chromosome pairs; an experienced geneticist can identify each band. In addition to the banding patterns, chromosomes are further identified on the basis of size and centromere location. To obtain the classic depiction of the karyotype in which homologous pairs of chromosomes are aligned in numerical order from longest to shortest, the geneticist obtains a digital image, identifies each chromosome, and manually arranges the chromosomes into this pattern ([\[link\]](#)).

At its most basic, the karyogram may reveal genetic abnormalities in which an individual has too many or too few chromosomes per cell. Examples of this are Down Syndrome, which is identified by a third copy of chromosome 21, and Turner Syndrome, which is characterized by the presence of only one X chromosome in women instead of the normal two. Geneticists can also identify large deletions or insertions of DNA. For instance, Jacobsen Syndrome—which involves distinctive facial features as well as heart and bleeding defects—is identified by a deletion on chromosome 11. Finally, the karyotype can pinpoint **translocations**, which occur when a segment of genetic material breaks from one chromosome and reattaches to another chromosome or to a different part of the same chromosome. Translocations are implicated in certain cancers, including chronic myelogenous leukemia.

During Mendel's lifetime, inheritance was an abstract concept that could only be inferred by performing crosses and observing the traits expressed by offspring. By observing a karyogram, today's geneticists can actually visualize the chromosomal composition of an individual to confirm or predict genetic abnormalities in offspring, even before birth.

## Disorders in Chromosome Number

Of all of the chromosomal disorders, abnormalities in chromosome number are the most obviously identifiable from a karyogram. Disorders of chromosome number include the duplication or loss of entire chromosomes, as well as changes in the number of complete sets of chromosomes. They are caused by **nondisjunction**, which occurs when pairs of homologous chromosomes or sister chromatids fail to separate during meiosis.

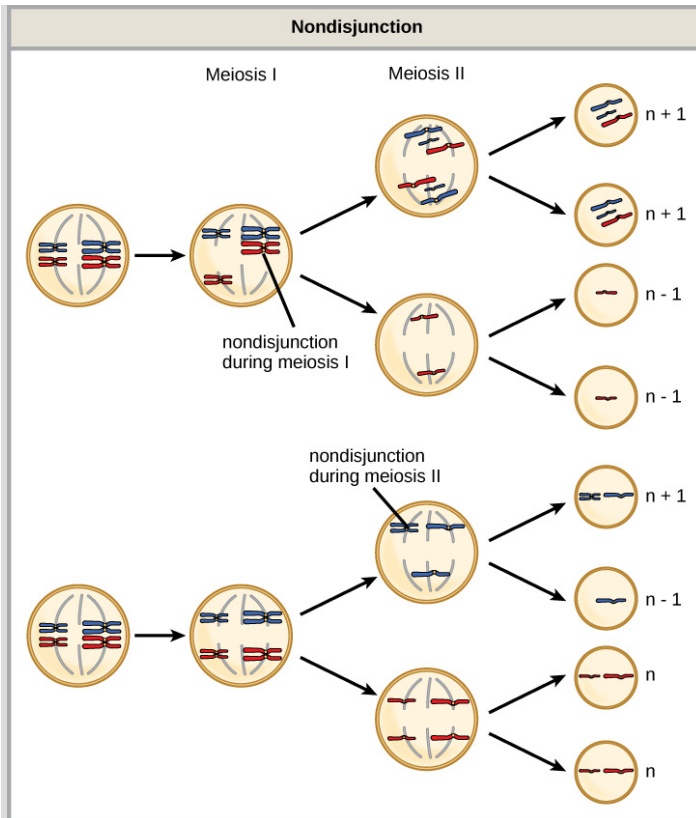
Misaligned or incomplete synapsis, or a dysfunction of the spindle apparatus that facilitates chromosome migration, can cause nondisjunction. The risk of nondisjunction occurring increases with the age of the parents.

Nondisjunction can occur during either meiosis I or II, with differing results ([link](#)). If homologous chromosomes fail to separate during meiosis I, the result is two gametes that lack that particular chromosome and two gametes with two copies of the chromosome. If sister chromatids fail to separate during meiosis II, the result is one gamete that lacks that chromosome, two normal gametes with one copy of the chromosome, and one gamete with two copies of the chromosome.

### **Note:**

Art Connection





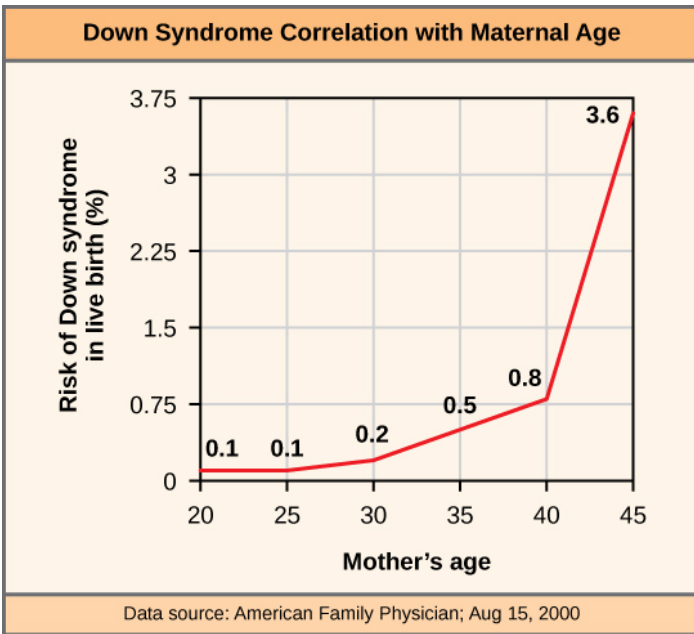
Nondisjunction occurs when homologous chromosomes or sister chromatids fail to separate during meiosis, resulting in an abnormal chromosome number. Nondisjunction may occur during meiosis I or meiosis II.

Which of the following statements about nondisjunction is true?

- Nondisjunction only results in gametes with  $n + 1$  or  $n - 1$  chromosomes.
- Nondisjunction occurring during meiosis II results in 50 percent normal gametes.
- Nondisjunction during meiosis I results in 50 percent normal gametes.
- Nondisjunction always results in four different kinds of gametes.

## Aneuploidy

An individual with the appropriate number of chromosomes for their species is called **euploid**; in humans, euploidy corresponds to 22 pairs of autosomes and one pair of sex chromosomes. An individual with an error in chromosome number is described as **aneuploid**, a term that includes **monosomy** (loss of one chromosome) or **trisomy** (gain of an extraneous chromosome). Monosomic human zygotes missing any one copy of an autosome invariably fail to develop to birth because they lack essential genes. This underscores the importance of “gene dosage” in humans. Most autosomal trisomies also fail to develop to birth; however, duplications of some of the smaller chromosomes (13, 15, 18, 21, or 22) can result in offspring that survive for several weeks to many years. Trisomic individuals suffer from a different type of genetic imbalance: an excess in gene dose. Individuals with an extra chromosome may synthesize an abundance of the gene products encoded by that chromosome. This extra dose (150 percent) of specific genes can lead to a number of functional challenges and often precludes development. The most common trisomy among viable births is that of chromosome 21, which corresponds to Down Syndrome. Individuals with this inherited disorder are characterized by short stature and stunted digits, facial distinctions that include a broad skull and large tongue, and significant developmental delays. The incidence of Down syndrome is correlated with maternal age; older women are more likely to become pregnant with fetuses carrying the trisomy 21 genotype ([\[link\]](#)).



The incidence of having a fetus with trisomy 21 increases dramatically with maternal age.

**Note:**

Link to Learning



Visualize the addition of a chromosome that leads to Down syndrome in this [video simulation](#).

## Polyploidy

An individual with more than the correct number of chromosome sets (two for diploid species) is called **polyploid**. For instance, fertilization of an abnormal diploid egg with a normal haploid sperm would yield a triploid zygote. Polyploid animals are extremely rare, with only a few examples among the flatworms, crustaceans, amphibians, fish, and lizards. Polyploid animals are sterile because meiosis cannot proceed normally and instead produces mostly aneuploid daughter cells that cannot yield viable zygotes. Rarely, polyploid animals can reproduce asexually by haplodiploidy, in which an unfertilized egg divides mitotically to produce offspring. In contrast, polyploidy is very common in the plant kingdom, and polyploid plants tend to be larger and more robust than euploids of their species ([link](#)).



As with many polyploid plants, this triploid orange daylily (*Hemerocallis fulva*) is particularly large and robust, and grows flowers with triple the number of petals of its diploid counterparts. (credit: Steve Karg)

## Sex Chromosome Nondisjunction in Humans

Humans display dramatic deleterious effects with autosomal trisomies and monosomies. Therefore, it may seem counterintuitive that human females and males can function normally, despite carrying different numbers of the X chromosome. Rather than a gain or loss of autosomes, variations in the number of sex chromosomes are associated with relatively mild effects. In part, this occurs because of a molecular process called **X inactivation**. Early in development, when female mammalian embryos consist of just a few thousand cells (relative to trillions in the newborn), one X chromosome in each cell inactivates by tightly condensing into a quiescent (dormant) structure called a Barr body. The chance that an X chromosome (maternally or paternally derived) is inactivated in each cell is random, but once the inactivation occurs, all cells derived from that one will have the same inactive X chromosome or Barr body. By this process, females compensate for their double genetic dose of X chromosome. In so-called “tortoiseshell” cats, embryonic X inactivation is observed as color variegation ([\[link\]](#)). Females that are heterozygous for an X-linked coat color gene will express one of two different coat colors over different regions of their body, corresponding to whichever X chromosome is inactivated in the embryonic cell progenitor of that region.



In cats, the gene for

coat color is located on the X chromosome.

In the embryonic development of female cats, one of the two X chromosomes is randomly inactivated in each cell, resulting in a tortoiseshell pattern if the cat has two different alleles for coat color. Male cats, having only one X chromosome, never exhibit a tortoiseshell coat color. (credit: Michael Bodega)

An individual carrying an abnormal number of X chromosomes will inactivate all but one X chromosome in each of her cells. However, even inactivated X chromosomes continue to express a few genes, and X chromosomes must reactivate for the proper maturation of female ovaries. As a result, X-chromosomal abnormalities are typically associated with mild mental and physical defects, as well as sterility. If the X chromosome is absent altogether, the individual will not develop in utero.

Several errors in sex chromosome number have been characterized. Individuals with three X chromosomes, called triplo-X, are phenotypically female but express developmental delays and reduced fertility. The XXY genotype, corresponding to one type of Klinefelter syndrome, corresponds to phenotypically male individuals with small testes, enlarged breasts, and reduced body hair. More complex types of Klinefelter syndrome exist in which the individual has as many as five X chromosomes. In all types, every X chromosome except one undergoes inactivation to compensate for the excess genetic dosage. This can be seen as several Barr bodies in each

cell nucleus. Turner syndrome, characterized as an XO genotype (i.e., only a single sex chromosome), corresponds to a phenotypically female individual with short stature, webbed skin in the neck region, hearing and cardiac impairments, and sterility.

## **Duplications and Deletions**

In addition to the loss or gain of an entire chromosome, a chromosomal segment may be duplicated or lost. Duplications and deletions often produce offspring that survive but exhibit physical and mental abnormalities. Duplicated chromosomal segments may fuse to existing chromosomes or may be free in the nucleus. Cri-du-chat (from the French for “cry of the cat”) is a syndrome associated with nervous system abnormalities and identifiable physical features that result from a deletion of most of 5p (the small arm of chromosome 5) ([link](#)). Infants with this genotype emit a characteristic high-pitched cry on which the disorder’s name is based.



This individual with cri-du-chat syndrome is shown at two, four, nine, and 12 years of age. (credit: Paola Cerruti Mainardi)

## Chromosomal Structural Rearrangements

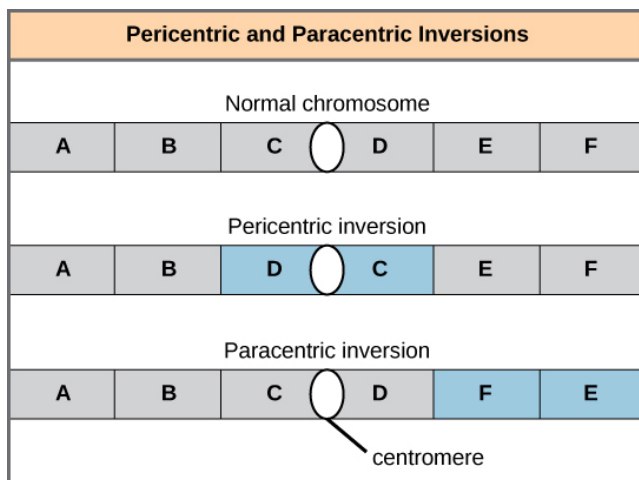
Cytologists have characterized numerous structural rearrangements in chromosomes, but chromosome inversions and translocations are the most common. Both are identified during meiosis by the adaptive pairing of rearranged chromosomes with their former homologs to maintain appropriate gene alignment. If the genes carried on two homologs are not oriented correctly, a recombination event could result in the loss of genes from one chromosome and the gain of genes on the other. This would produce aneuploid gametes.



## Chromosome Inversions

A **chromosome inversion** is the detachment, 180° rotation, and reinsertion of part of a chromosome. Inversions may occur in nature as a result of mechanical shear, or from the action of transposable elements (special DNA sequences capable of facilitating the rearrangement of chromosome segments with the help of enzymes that cut and paste DNA sequences). Unless they disrupt a gene sequence, inversions only change the orientation of genes and are likely to have more mild effects than aneuploid errors. However, altered gene orientation can result in functional changes because regulators of gene expression could be moved out of position with respect to their targets, causing aberrant levels of gene products.

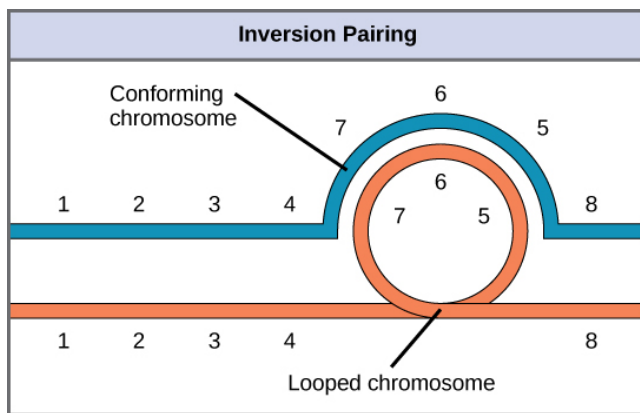
An inversion can be **pericentric** and include the centromere, or **paracentric** and occur outside of the centromere ([\[link\]](#)). A pericentric inversion that is asymmetric about the centromere can change the relative lengths of the chromosome arms, making these inversions easily identifiable.



Pericentric inversions include the centromere, and paracentric inversions do not. A pericentric inversion can change the relative

lengths of the chromosome arms;  
a paracentric inversion cannot.

When one homologous chromosome undergoes an inversion but the other does not, the individual is described as an inversion heterozygote. To maintain point-for-point synapsis during meiosis, one homolog must form a loop, and the other homolog must mold around it. Although this topology can ensure that the genes are correctly aligned, it also forces the homologs to stretch and can be associated with regions of imprecise synapsis ([\[link\]](#)).



When one chromosome undergoes an inversion but the other does not, one chromosome must form an inverted loop to retain point-for-point interaction during synapsis. This inversion pairing is essential to maintaining gene alignment during meiosis and to allow for recombination.

**Note:**

## Evolution Connection

### The Chromosome 18 Inversion

Not all structural rearrangements of chromosomes produce nonviable, impaired, or infertile individuals. In rare instances, such a change can result in the evolution of a new species. In fact, a pericentric inversion in chromosome 18 appears to have contributed to the evolution of humans. This inversion is not present in our closest genetic relatives, the chimpanzees. Humans and chimpanzees differ cytogenetically by pericentric inversions on several chromosomes and by the fusion of two separate chromosomes in chimpanzees that correspond to chromosome two in humans.

The pericentric chromosome 18 inversion is believed to have occurred in early humans following their divergence from a common ancestor with chimpanzees approximately five million years ago. Researchers characterizing this inversion have suggested that approximately 19,000 nucleotide bases were duplicated on 18p, and the duplicated region inverted and reinserted on chromosome 18 of an ancestral human.

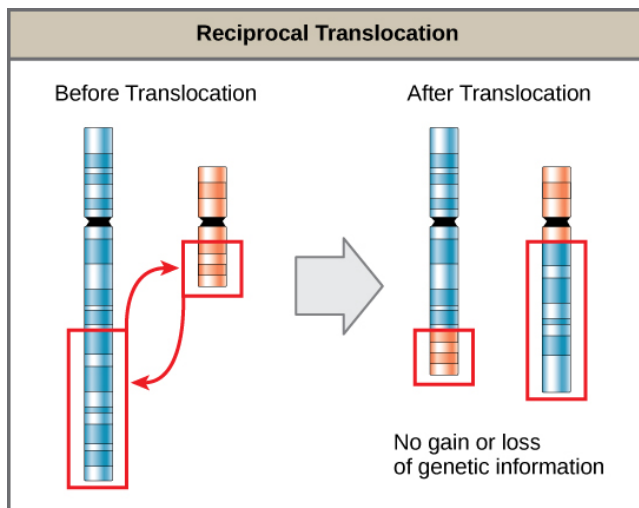
A comparison of human and chimpanzee genes in the region of this inversion indicates that two genes—*ROCK1* and *USP14*—that are adjacent on chimpanzee chromosome 17 (which corresponds to human chromosome 18) are more distantly positioned on human chromosome 18. This suggests that one of the inversion breakpoints occurred between these two genes.

Interestingly, humans and chimpanzees express *USP14* at distinct levels in specific cell types, including cortical cells and fibroblasts. Perhaps the chromosome 18 inversion in an ancestral human repositioned specific genes and reset their expression levels in a useful way. Because both *ROCK1* and *USP14* encode cellular enzymes, a change in their expression could alter cellular function. It is not known how this inversion contributed to hominid evolution, but it appears to be a significant factor in the divergence of humans from other primates. [\[footnote\]](#)

Violaine Goidts et al., “Segmental duplication associated with the human-specific inversion of chromosome 18: a further example of the impact of segmental duplications on karyotype and genome evolution in primates,” *Human Genetics*. 115 (2004):116-122

## Translocations

A **translocation** occurs when a segment of a chromosome dissociates and reattaches to a different, nonhomologous chromosome. Translocations can be benign or have devastating effects depending on how the positions of genes are altered with respect to regulatory sequences. Notably, specific translocations have been associated with several cancers and with schizophrenia. Reciprocal translocations result from the exchange of chromosome segments between two nonhomologous chromosomes such that there is no gain or loss of genetic information ([link](#)).



A reciprocal translocation occurs when a segment of DNA is transferred from one chromosome to another, nonhomologous chromosome. (credit: modification of work by National Human Genome Research/USA)

## Section Summary

The number, size, shape, and banding pattern of chromosomes make them easily identifiable in a karyogram and allows for the assessment of many chromosomal abnormalities. Disorders in chromosome number, or aneuploidies, are typically lethal to the embryo, although a few trisomic genotypes are viable. Because of X inactivation, aberrations in sex chromosomes typically have milder phenotypic effects. Aneuploidies also include instances in which segments of a chromosome are duplicated or deleted. Chromosome structures may also be rearranged, for example by inversion or translocation. Both of these aberrations can result in problematic phenotypic effects. Because they force chromosomes to assume unnatural topologies during meiosis, inversions and translocations are often associated with reduced fertility because of the likelihood of nondisjunction.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Which of the following statements about nondisjunction is true?

- a. Nondisjunction only results in gametes with  $n+1$  or  $n-1$  chromosomes.
- b. Nondisjunction occurring during meiosis II results in 50 percent normal gametes.
- c. Nondisjunction during meiosis I results in 50 percent normal gametes.
- d. Nondisjunction always results in four different kinds of gametes.

---

#### Solution:

[\[link\]](#) B.

## Review Questions

**Exercise:**

**Problem:**

Which of the following codes describes position 12 on the long arm of chromosome 13?

- a. 13p12
- b. 13q12
- c. 12p13
- d. 12q13

---

**Solution:**

B

**Exercise:**

**Problem:**

In agriculture, polyploid crops (like coffee, strawberries, or bananas) tend to produce \_\_\_\_\_.

- a. more uniformity
- b. more variety
- c. larger yields
- d. smaller yields

---

**Solution:**

C

**Exercise:**

**Problem:**

Assume a pericentric inversion occurred in one of two homologs prior to meiosis. The other homolog remains normal. During meiosis, what structure—if any—would these homologs assume in order to pair accurately along their lengths?

- a. V formation
- b. cruciform
- c. loop
- d. pairing would not be possible

---

**Solution:**

C

**Exercise:**

**Problem:** The genotype XXY corresponds to

- a. Klinefelter syndrome
- b. Turner syndrome
- c. Triplo-X
- d. Jacob syndrome

---

**Solution:**

A

**Exercise:****Problem:**

Abnormalities in the number of X chromosomes tends to have milder phenotypic effects than the same abnormalities in autosomes because of \_\_\_\_\_.

- a. deletions
- b. nonhomologous recombination
- c. synapsis
- d. X inactivation

---

**Solution:**

D

**Exercise:**

**Problem:** By definition, a pericentric inversion includes the \_\_\_\_\_.

- a. centromere
- b. chiasma
- c. telomere
- d. synapse

---

**Solution:**

A

## Free Response

**Exercise:**

**Problem:**

Using diagrams, illustrate how nondisjunction can result in an aneuploid zygote.

---

**Solution:**

Exact diagram style will vary; diagram should look like [\[link\]](#).



## Glossary

### aneuploid

individual with an error in chromosome number; includes deletions and duplications of chromosome segments

### autosome

any of the non-sex chromosomes

### chromosome inversion

detachment, 180° rotation, and reinsertion of a chromosome arm

### euploid

individual with the appropriate number of chromosomes for their species

### karyogram

photographic image of a karyotype

### karyotype

number and appearance of an individual's chromosomes; includes the size, banding patterns, and centromere position

### monosomy

otherwise diploid genotype in which one chromosome is missing

### nondisjunction

failure of synapsed homologs to completely separate and migrate to separate poles during the first cell division of meiosis

### paracentric

inversion that occurs outside of the centromere

### pericentric

inversion that involves the centromere

### polyploid

individual with an incorrect number of chromosome sets

translocation

process by which one segment of a chromosome dissociates and reattaches to a different, nonhomologous chromosome

trisomy

otherwise diploid genotype in which one entire chromosome is duplicated

X inactivation

condensation of X chromosomes into Barr bodies during embryonic development in females to compensate for the double genetic dose

## Introduction

class="introduction"

Dolly  
the  
sheep  
was the  
first  
large  
mamma  
l to be  
cloned.



The three letters “DNA” have now become synonymous with crime solving, paternity testing, human identification, and genetic testing. DNA can be retrieved from hair, blood, or saliva. Each person’s DNA is unique, and it is possible to detect differences between individuals within a species on the basis of these unique features.

DNA analysis has many practical applications beyond forensics. In humans, DNA testing is applied to numerous uses: determining paternity, tracing

genealogy, identifying pathogens, archeological research, tracing disease outbreaks, and studying human migration patterns. In the medical field, DNA is used in diagnostics, new vaccine development, and cancer therapy. It is now possible to determine predisposition to diseases by looking at genes.

Each human cell has 23 pairs of chromosomes: one set of chromosomes is inherited from the mother and the other set is inherited from the father. There is also a mitochondrial genome, inherited exclusively from the mother, which can be involved in inherited genetic disorders. On each chromosome, there are thousands of genes that are responsible for determining the genotype and phenotype of the individual. A gene is defined as a sequence of DNA that codes for a functional product. The human haploid genome contains 3 billion base pairs and has between 20,000 and 25,000 functional genes.

## Historical Basis of Modern Understanding

By the end of this section, you will be able to:

- Explain transformation of DNA
- Describe the key experiments that helped identify that DNA is the genetic material
- State and explain Chargaff's rules

Modern understandings of DNA have evolved from the discovery of nucleic acid to the development of the double-helix model. In the 1860s, Friedrich Miescher ([link](#)), a physician by profession, was the first person to isolate phosphate-rich chemicals from white blood cells or leukocytes. He named these chemicals (which would eventually be known as RNA and DNA) nuclein because they were isolated from the nuclei of the cells.



Friedrich  
Miescher  
(1844–1895)  
discovered  
nucleic acids.

### **Note:**

Link to Learning



To see Miescher conduct an experiment step-by-step, click through [this review](#) of how he discovered the key role of DNA and proteins in the nucleus.

A half century later, British bacteriologist Frederick Griffith was perhaps the first person to show that hereditary information could be transferred from one cell to another “horizontally,” rather than by descent. In 1928, he reported the first demonstration of bacterial **transformation**, a process in which external DNA is taken up by a cell, thereby changing morphology and physiology. He was working with *Streptococcus pneumoniae*, the bacterium that causes pneumonia. Griffith worked with two strains, rough (R) and smooth (S). The R strain is non-pathogenic (does not cause disease) and is called rough because its outer surface is a cell wall and lacks a capsule; as a result, the cell surface appears uneven under the microscope. The S strain is pathogenic (disease-causing) and has a capsule outside its cell wall. As a result, it has a smooth appearance under the microscope. Griffith injected the live R strain into mice and they survived. In another experiment, when he injected mice with the heat-killed S strain, they also survived. In a third set of experiments, a mixture of live R strain and heat-killed S strain were injected into mice, and—to his surprise—the mice died. Upon isolating the live bacteria from the dead mouse, only the S strain of bacteria was recovered. When this isolated S strain was injected into fresh mice, the mice died. Griffith concluded that something had passed from the heat-killed S strain into the live R strain and transformed it into the pathogenic S strain, and he called this the transforming principle ([\[link\]](#)). These experiments are now famously known as Griffith's transformation experiments.



Mouse injected with heat-killed virulent S strain lives.



Mouse injected with both heat-killed S strain and live non-virulent R strain dies.

Two strains of *S. pneumoniae* were used in Griffith's transformation experiments. The R strain is non-pathogenic. The S strain is pathogenic and causes death. When Griffith injected a mouse with the heat-killed S strain and a live R strain, the mouse died. The S strain was recovered from the dead mouse. Thus, Griffith concluded that something had passed from the heat-killed S strain to the R strain, transforming the R strain into S strain in the process. (credit "living mouse": modification of work by NIH; credit "dead mouse": modification of work by Sarah Marriage)

Scientists Oswald Avery, Colin MacLeod, and Maclyn McCarty (1944) were interested in exploring this transforming principle further. They isolated the S strain from the dead mice and isolated the proteins and nucleic acids, namely RNA and DNA, as these were possible candidates for the molecule of heredity. They conducted a systematic elimination study. They used enzymes that specifically degraded each component and then used each mixture separately to transform the R strain. They found that when DNA was degraded, the resulting mixture was no longer able to transform the bacteria, whereas all of the other combinations were able to transform the bacteria. This led them to conclude that DNA was the transforming principle.

**Note:****Career Connection****Forensic Scientists and DNA Analysis**

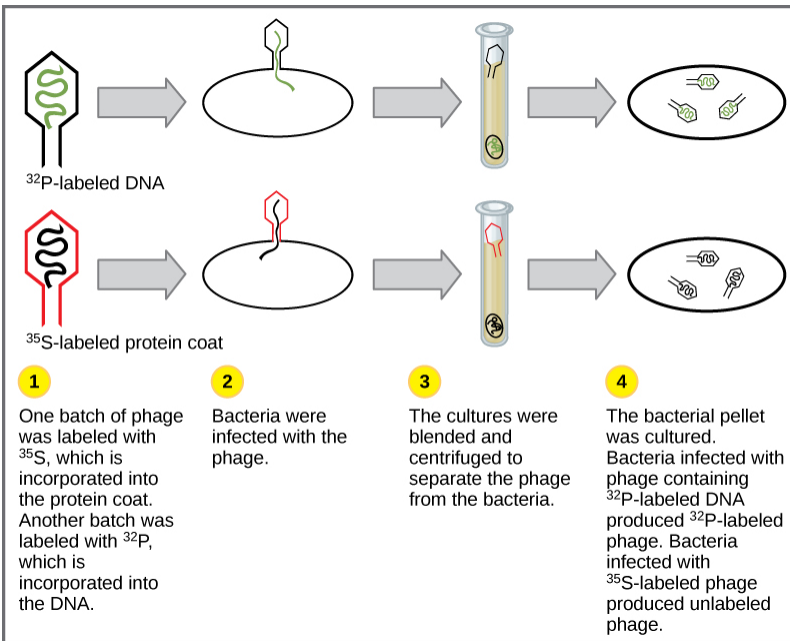
DNA evidence was used for the first time to solve an immigration case. The story started with a teenage boy returning to London from Ghana to be with his mother. Immigration authorities at the airport were suspicious of him, thinking that he was traveling on a forged passport. After much persuasion, he was allowed to go live with his mother, but the immigration authorities did not drop the case against him. All types of evidence, including photographs, were provided to the authorities, but deportation proceedings were started nevertheless. Around the same time, Dr. Alec Jeffreys of Leicester University in the United Kingdom had invented a technique known as DNA fingerprinting. The immigration authorities approached Dr. Jeffreys for help. He took DNA samples from the mother and three of her children, plus an unrelated mother, and compared the samples with the boy's DNA. Because the biological father was not in the picture, DNA from the three children was compared with the boy's DNA. He found a match in the boy's DNA for both the mother and his three siblings. He concluded that the boy was indeed the mother's son.

Forensic scientists analyze many items, including documents, handwriting, firearms, and biological samples. They analyze the DNA content of hair, semen, saliva, and blood, and compare it with a database of DNA profiles of known criminals. Analysis includes DNA isolation, sequencing, and sequence analysis; most forensic DNA analysis involves polymerase chain reaction (PCR) amplification of short tandem repeat (STR) loci and electrophoresis to determine the length of the PCR-amplified fragment. Only mitochondrial DNA is sequenced for forensics. Forensic scientists are expected to appear at court hearings to present their findings. They are usually employed in crime labs of city and state government agencies. Geneticists experimenting with DNA techniques also work for scientific and research organizations, pharmaceutical industries, and college and university labs. Students wishing to pursue a career as a forensic scientist should have at least a bachelor's degree in chemistry, biology, or physics, and preferably some experience working in a laboratory.



Experiments conducted by Martha Chase and Alfred Hershey in 1952 provided confirmatory evidence that DNA was the genetic material and not proteins. Chase and Hershey were studying a bacteriophage, which is a virus that infects bacteria. Viruses typically have a simple structure: a protein coat, called the capsid, and a nucleic acid core that contains the genetic material, either DNA or RNA. The bacteriophage infects the host bacterial cell by attaching to its surface, and then it injects its nucleic acids inside the cell. The phage DNA makes multiple copies of itself using the host machinery, and eventually the host cell bursts, releasing a large number of bacteriophages. Hershey and Chase labeled one batch of phage with radioactive sulfur,  $^{35}\text{S}$ , to label the protein coat. Another batch of phage were labeled with radioactive phosphorus,  $^{32}\text{P}$ . Because phosphorus is found in DNA, but not protein, the DNA and not the protein would be tagged with radioactive phosphorus.

Each batch of phage was allowed to infect the cells separately. After infection, the phage bacterial suspension was put in a blender, which caused the phage coat to be detached from the host cell. The phage and bacterial suspension was spun down in a centrifuge. The heavier bacterial cells settled down and formed a pellet, whereas the lighter phage particles stayed in the supernatant. In the tube that contained phage labeled with  $^{35}\text{S}$ , the supernatant contained the radioactively labeled phage, whereas no radioactivity was detected in the pellet. In the tube that contained the phage labeled with  $^{32}\text{P}$ , the radioactivity was detected in the pellet that contained the heavier bacterial cells, and no radioactivity was detected in the supernatant. Hershey and Chase concluded that it was the phage DNA that was injected into the cell and carried information to produce more phage particles, thus providing evidence that DNA was the genetic material and not proteins ([\[link\]](#)).



In Hershey and Chase's experiments, bacteria were infected with phage radiolabeled with either  $^{35}\text{S}$ , which labels protein, or  $^{32}\text{P}$ , which labels DNA. Only  $^{32}\text{P}$  entered the bacterial cells, indicating that DNA is the genetic material.

Around this same time, Austrian biochemist Erwin Chargaff examined the content of DNA in different species and found that the amounts of adenine, thymine, guanine, and cytosine were not found in equal quantities, and that it varied from species to species, but not between individuals of the same species. He found that the amount of adenine equals the amount of thymine, and the amount of cytosine equals the amount of guanine, or  $A = T$  and  $G = C$ . This is also known as Chargaff's rules. This finding proved immensely useful when Watson and Crick were getting ready to propose their DNA double helix model.

## Section Summary

DNA was first isolated from white blood cells by Friedrich Miescher, who called it nuclein because it was isolated from nuclei. Frederick Griffith's experiments with strains of *Streptococcus pneumoniae* provided the first hint that DNA may be the transforming principle. Avery, MacLeod, and McCarty proved that DNA is required for the transformation of bacteria. Later experiments by Hershey and Chase using bacteriophage T2 proved that DNA is the genetic material. Chargaff found that the ratio of A = T and C = G, and that the percentage content of A, T, G, and C is different for different species.

## Review Questions

### Exercise:

#### Problem:

If DNA of a particular species was analyzed and it was found that it contains 27 percent A, what would be the percentage of C?

- a. 27 percent
- b. 30 percent
- c. 23 percent
- d. 54 percent

---

#### Solution:

C

### Exercise:

#### Problem:

The experiments by Hershey and Chase helped confirm that DNA was the hereditary material on the basis of the finding that:

- a. radioactive phage were found in the pellet
- b. radioactive cells were found in the supernatant
- c. radioactive sulfur was found inside the cell

d. radioactive phosphorus was found in the cell

---

**Solution:**

D

**Free Response**

**Exercise:**

**Problem:**

Explain Griffith's transformation experiments. What did he conclude from them?

---

**Solution:**

Live R cells acquired genetic information from the heat-killed S cells that “transformed” the R cells into S cells.

**Exercise:**

**Problem:**

Why were radioactive sulfur and phosphorous used to label bacteriophage in Hershey and Chase's experiments?

---

**Solution:**

Sulfur is an element found in proteins and phosphorus is a component of nucleic acids.

**Glossary**

transformation

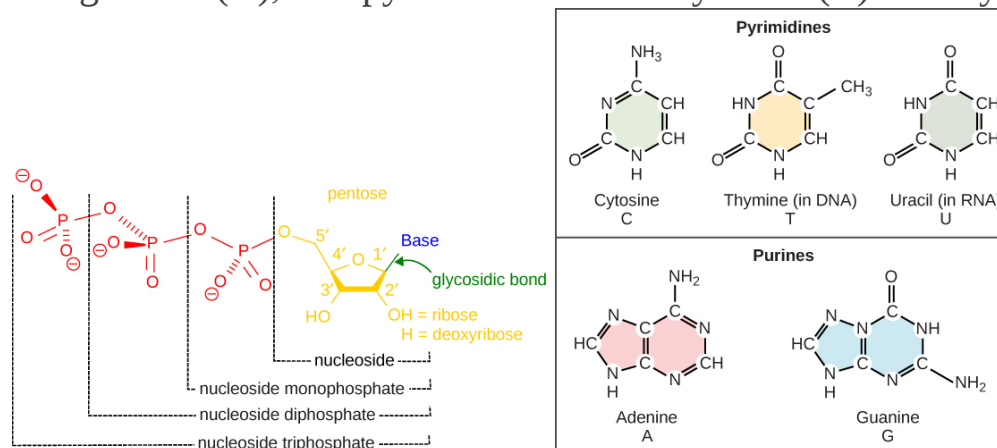
process in which external DNA is taken up by a cell

## DNA Structure and Sequencing

By the end of this section, you will be able to:

- Describe the structure of DNA
- Explain the Sanger method of DNA sequencing
- Discuss the similarities and differences between eukaryotic and prokaryotic DNA

The building blocks of DNA are nucleotides. The important components of the nucleotide are a nitrogenous base, deoxyribose (5-carbon sugar), and a phosphate group ([\[link\]](#)). The nucleotide is named depending on the nitrogenous base. The nitrogenous base can be a purine such as adenine (A) and guanine (G), or a pyrimidine such as cytosine (C) and thymine (T).



Each nucleotide is made up of a sugar, a phosphate group, and a nitrogenous base. The sugar is deoxyribose in DNA and ribose in RNA.

The nucleotides combine with each other by covalent bonds known as phosphodiester bonds or linkages. The purines have a double ring structure with a six-membered ring fused to a five-membered ring. Pyrimidines are smaller in size; they have a single six-membered ring structure. The carbon atoms of the five-carbon sugar are numbered 1', 2', 3', 4', and 5' (1' is read as "one prime"). The phosphate residue is attached to the hydroxyl group of the 5' carbon of one sugar of one nucleotide and the hydroxyl group of the

3' carbon of the sugar of the next nucleotide, thereby forming a 5'-3' phosphodiester bond.

In the 1950s, Francis Crick and James Watson worked together to determine the structure of DNA at the University of Cambridge, England. Other scientists like Linus Pauling and Maurice Wilkins were also actively exploring this field. Pauling had discovered the secondary structure of proteins using X-ray crystallography. In Wilkins' lab, researcher Rosalind Franklin was using X-ray diffraction methods to understand the structure of DNA. Watson and Crick were able to piece together the puzzle of the DNA molecule on the basis of Franklin's data because Crick had also studied X-ray diffraction ([\[link\]](#)). In 1962, James Watson, Francis Crick, and Maurice Wilkins were awarded the Nobel Prize in Medicine. Unfortunately, by then Franklin had died, and Nobel prizes are not awarded posthumously.



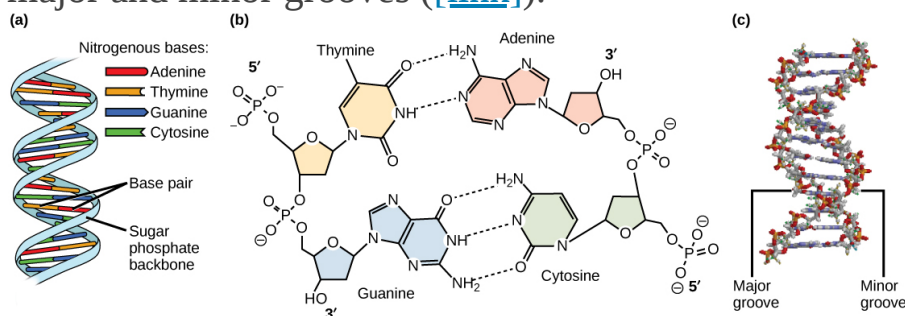
(a)



(b)

The work of pioneering scientists (a) James Watson, Francis Crick, and Maclyn McCarty led to our present day understanding of DNA. Scientist Rosalind Franklin discovered (b) the X-ray diffraction pattern of DNA, which helped to elucidate its double helix structure. (credit a: modification of work by Marjorie McCarty, Public Library of Science)

Watson and Crick proposed that DNA is made up of two strands that are twisted around each other to form a right-handed helix. Base pairing takes place between a purine and pyrimidine; namely, A pairs with T and G pairs with C. Adenine and thymine are complementary base pairs, and cytosine and guanine are also complementary base pairs. The base pairs are stabilized by hydrogen bonds; adenine and thymine form two hydrogen bonds and cytosine and guanine form three hydrogen bonds. The two strands are anti-parallel in nature; that is, the 3' end of one strand faces the 5' end of the other strand. The sugar and phosphate of the nucleotides form the backbone of the structure, whereas the nitrogenous bases are stacked inside. Each base pair is separated from the other base pair by a distance of 0.34 nm, and each turn of the helix measures 3.4 nm. Therefore, ten base pairs are present per turn of the helix. The diameter of the DNA double helix is 2 nm, and it is uniform throughout. Only the pairing between a purine and pyrimidine can explain the uniform diameter. The twisting of the two strands around each other results in the formation of uniformly spaced major and minor grooves ([\[link\]](#)).



DNA has (a) a double helix structure and (b) phosphodiester bonds. The (c) major and minor grooves are binding sites for DNA binding proteins during processes such as transcription (the copying of RNA from DNA) and replication.

## DNA Sequencing Techniques

Until the 1990s, the sequencing of DNA (reading the sequence of DNA) was a relatively expensive and long process. Using radiolabeled nucleotides

also compounded the problem through safety concerns. With currently available technology and automated machines, the process is cheap, safer, and can be completed in a matter of hours. Fred Sanger developed the sequencing method used for the human genome sequencing project, which is widely used today ([link](#)).

**Note:**

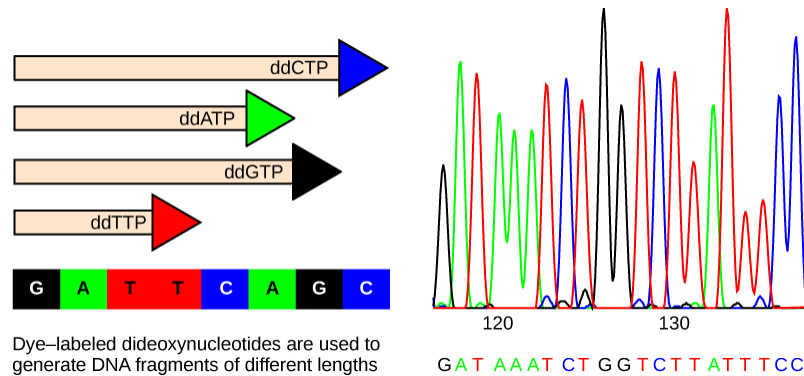
Link to Learning



Visit [this site](#) to watch a video explaining the DNA sequence reading technique that resulted from Sanger's work.

The method is known as the dideoxy chain termination method. The sequencing method is based on the use of chain terminators, the dideoxynucleotides (ddNTPs). The dideoxynucleotides, or ddNTPs, differ from the deoxynucleotides by the lack of a free 3' OH group on the five-carbon sugar. If a ddNTP is added to a growing a DNA strand, the chain is not extended any further because the free 3' OH group needed to add another nucleotide is not available. By using a predetermined ratio of deoxyribonucleotides to dideoxynucleotides, it is possible to generate DNA fragments of different sizes.





In Frederick Sanger's dideoxy chain termination method, dye-labeled dideoxynucleotides are used to generate DNA fragments that terminate at different points. The DNA is separated by capillary electrophoresis on the basis of size, and from the order of fragments formed, the DNA sequence can be read. The DNA sequence readout is shown on an electropherogram that is generated by a laser scanner.

The DNA sample to be sequenced is denatured or separated into two strands by heating it to high temperatures. The DNA is divided into four tubes in which a primer, DNA polymerase, and all four nucleotides (A, T, G, and C) are added. In addition to each of the four tubes, limited quantities of one of the four dideoxynucleotides are added to each tube respectively. The tubes are labeled as A, T, G, and C according to the ddNTP added. For detection purposes, each of the four dideoxynucleotides carries a different fluorescent label. Chain elongation continues until a fluorescent dideoxy nucleotide is incorporated, after which no further elongation takes place. After the reaction is over, electrophoresis is performed. Even a difference in length of a single base can be detected. The sequence is read from a laser scanner. For his work on DNA sequencing, Sanger received a Nobel Prize in chemistry in 1980.

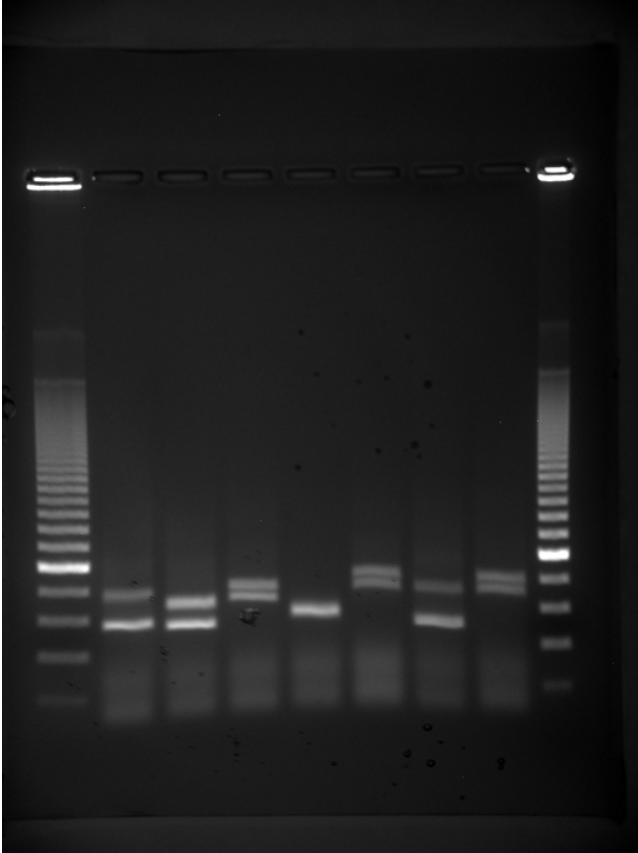
**Note:**

## Link to Learning



Sanger's genome sequencing has led to a race to sequence human genomes at a rapid speed and low cost, often referred to as the \$1000 in one day sequence. Learn more by selecting the Sequencing at Speed animation [here](#).

Gel **electrophoresis** is a technique used to separate DNA fragments of different sizes. Usually the gel is made of a chemical called agarose. Agarose powder is added to a buffer and heated. After cooling, the gel solution is poured into a casting tray. Once the gel has solidified, the DNA is loaded on the gel and electric current is applied. The DNA has a net negative charge and moves from the negative electrode toward the positive electrode. The electric current is applied for sufficient time to let the DNA separate according to size; the smallest fragments will be farthest from the well (where the DNA was loaded), and the heavier molecular weight fragments will be closest to the well. Once the DNA is separated, the gel is stained with a DNA-specific dye for viewing it ([link](#)).



DNA can be separated on the basis of size using gel electrophoresis.  
(credit: James Jacob, Tompkins Cortland Community College)

**Note:**

Evolution Connection

**Neanderthal Genome: How Are We Related?**

The first draft sequence of the Neanderthal genome was recently published by Richard E. Green et al. in 2010. [\[footnote\]](#) Neanderthals are the closest ancestors of present-day humans. They were known to have lived in Europe and Western Asia before they disappeared from fossil records approximately 30,000 years ago. Green's team studied almost 40,000-year-old fossil remains that were selected from sites across the world. Extremely

sophisticated means of sample preparation and DNA sequencing were employed because of the fragile nature of the bones and heavy microbial contamination. In their study, the scientists were able to sequence some four billion base pairs. The Neanderthal sequence was compared with that of present-day humans from across the world. After comparing the sequences, the researchers found that the Neanderthal genome had 2 to 3 percent greater similarity to people living outside Africa than to people in Africa. While current theories have suggested that all present-day humans can be traced to a small ancestral population in Africa, the data from the Neanderthal genome may contradict this view. Green and his colleagues also discovered DNA segments among people in Europe and Asia that are more similar to Neanderthal sequences than to other contemporary human sequences. Another interesting observation was that Neanderthals are as closely related to people from Papua New Guinea as to those from China or France. This is surprising because Neanderthal fossil remains have been located only in Europe and West Asia. Most likely, genetic exchange took place between Neanderthals and modern humans as modern humans emerged out of Africa, before the divergence of Europeans, East Asians, and Papua New Guineans.

Richard E. Green et al., “A Draft Sequence of the Neandertal Genome,” *Science* 328 (2010): 710-22.

Several genes seem to have undergone changes from Neanderthals during the evolution of present-day humans. These genes are involved in cranial structure, metabolism, skin morphology, and cognitive development. One of the genes that is of particular interest is *RUNX2*, which is different in modern day humans and Neanderthals. This gene is responsible for the prominent frontal bone, bell-shaped rib cage, and dental differences seen in Neanderthals. It is speculated that an evolutionary change in *RUNX2* was important in the origin of modern-day humans, and this affected the cranium and the upper body.

**Note:**

[Link to Learning](#)



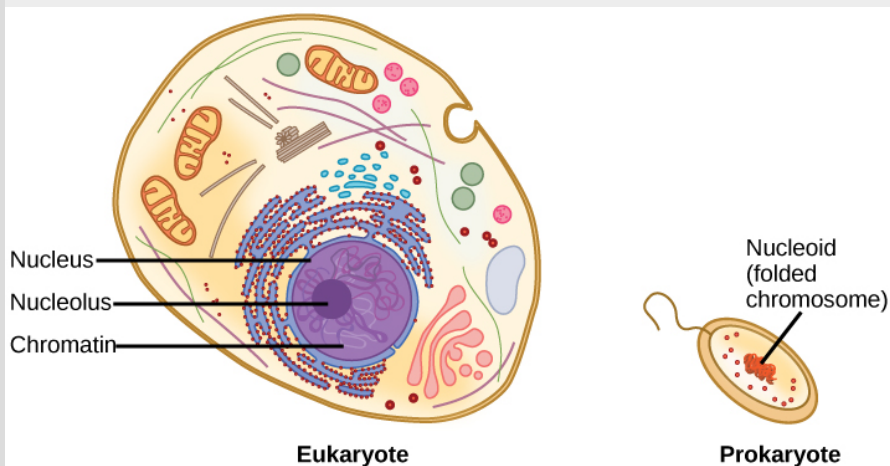
Watch [Svante Pääbo's talk](#) explaining the Neanderthal genome research at the 2011 annual TED (Technology, Entertainment, Design) conference.

## DNA Packaging in Cells

When comparing prokaryotic cells to eukaryotic cells, prokaryotes are much simpler than eukaryotes in many of their features ([link](#)). Most prokaryotes contain a single, circular chromosome that is found in an area of the cytoplasm called the nucleoid.

### Note:

#### Art Connection



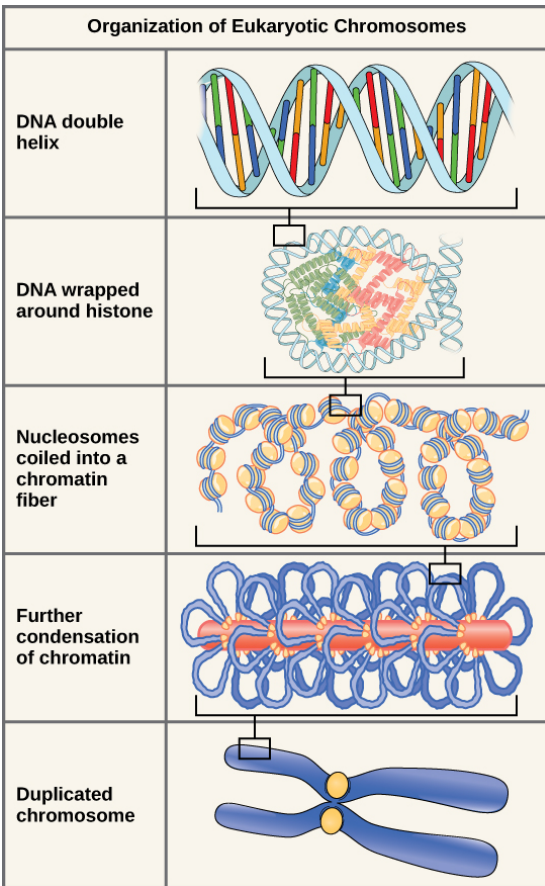
A eukaryote contains a well-defined nucleus, whereas in prokaryotes, the chromosome lies in the cytoplasm in an area called the nucleoid.

In eukaryotic cells, DNA and RNA synthesis occur in a separate compartment from protein synthesis. In prokaryotic cells, both processes occur together. What advantages might there be to separating the processes? What advantages might there be to having them occur together?

The size of the genome in one of the most well-studied prokaryotes, *E.coli*, is 4.6 million base pairs (approximately 1.1 mm, if cut and stretched out). So how does this fit inside a small bacterial cell? The DNA is twisted by what is known as supercoiling. Supercoiling means that DNA is either under-wound (less than one turn of the helix per 10 base pairs) or over-wound (more than 1 turn per 10 base pairs) from its normal relaxed state. Some proteins are known to be involved in the supercoiling; other proteins and enzymes such as DNA gyrase help in maintaining the supercoiled structure.

Eukaryotes, whose chromosomes each consist of a linear DNA molecule, employ a different type of packing strategy to fit their DNA inside the nucleus ([\[link\]](#)). At the most basic level, DNA is wrapped around proteins known as histones to form structures called nucleosomes. The histones are evolutionarily conserved proteins that are rich in basic amino acids and form an octamer. The DNA (which is negatively charged because of the phosphate groups) is wrapped tightly around the histone core. This nucleosome is linked to the next one with the help of a linker DNA. This is also known as the “beads on a string” structure. This is further compacted into a 30 nm fiber, which is the diameter of the structure. At the metaphase stage, the chromosomes are at their most compact, are approximately 700 nm in width, and are found in association with scaffold proteins.

In interphase, eukaryotic chromosomes have two distinct regions that can be distinguished by staining. The tightly packaged region is known as heterochromatin, and the less dense region is known as euchromatin. Heterochromatin usually contains genes that are not expressed, and is found in the regions of the centromere and telomeres. The euchromatin usually contains genes that are transcribed, with DNA packaged around nucleosomes but not further compacted.



These figures illustrate the compaction of the eukaryotic chromosome.

## Section Summary

The currently accepted model of the double-helix structure of DNA was proposed by Watson and Crick. Some of the salient features are that the two strands that make up the double helix are complementary and anti-parallel in nature. Deoxyribose sugars and phosphates form the backbone of the structure, and the nitrogenous bases are stacked inside. The diameter of the double helix, 2 nm, is uniform throughout. A purine always pairs with a pyrimidine; A pairs with T, and G pairs with C. One turn of the helix has ten base pairs. During cell division, each daughter cell receives a copy of

the DNA by a process known as DNA replication. Prokaryotes are much simpler than eukaryotes in many of their features. Most prokaryotes contain a single, circular chromosome. In general, eukaryotic chromosomes contain a linear DNA molecule packaged into nucleosomes, and have two distinct regions that can be distinguished by staining, reflecting different states of packaging and compaction.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) In eukaryotic cells, DNA and RNA synthesis occur in a separate compartment from protein synthesis. In prokaryotic cells, both processes occur together. What advantages might there be to separating the processes? What advantages might there be to having them occur together?

---

#### Solution:

[\[link\]](#) Compartmentalization enables a eukaryotic cell to divide processes into discrete steps so it can build more complex protein and RNA products. But there is an advantage to having a single compartment as well: RNA and protein synthesis occurs much more quickly in a prokaryotic cell.

## Review Questions

### Exercise:

**Problem:** DNA double helix does not have which of the following?

- a. antiparallel configuration
- b. complementary base pairing
- c. major and minor grooves
- d. uracil



---

**Solution:**

D

**Exercise:**

**Problem:**In eukaryotes, what is the DNA wrapped around?

- a. single-stranded binding proteins
- b. sliding clamp
- c. polymerase
- d. histones

---

**Solution:**

D

**Free Response**

**Exercise:**

**Problem:**Provide a brief summary of the Sanger sequencing method.

---

**Solution:**

The template DNA strand is mixed with a DNA polymerase, a primer, the 4 deoxynucleotides, and a limiting concentration of 4 dideoxynucleotides. DNA polymerase synthesizes a strand complementary to the template. Incorporation of ddNTPs at different locations results in DNA fragments that have terminated at every possible base in the template. These fragments are separated by gel electrophoresis and visualized by a laser detector to determine the sequence of bases.

**Exercise:**

**Problem:**

Describe the structure and complementary base pairing of DNA.

---

**Solution:**

DNA has two strands in anti-parallel orientation. The sugar-phosphate linkages form a backbone on the outside, and the bases are paired on the inside: A with T, and G with C, like rungs on a spiral ladder.

**Glossary**

electrophoresis

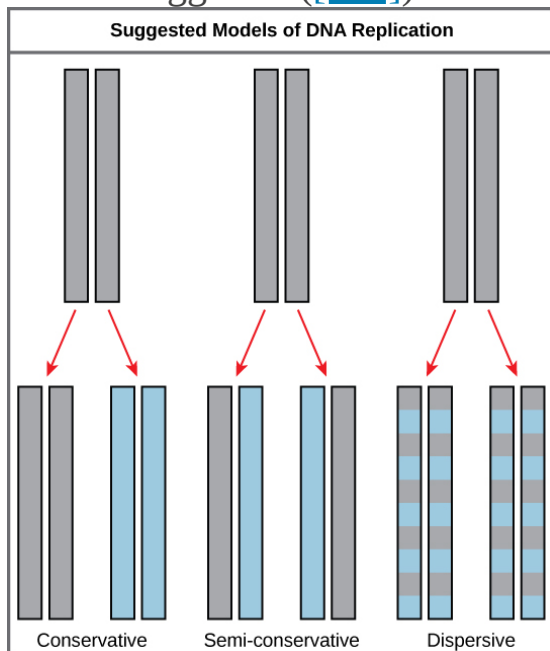
technique used to separate DNA fragments according to size

## Basics of DNA Replication

By the end of this section, you will be able to:

- Explain how the structure of DNA reveals the replication process
- Describe the Meselson and Stahl experiments

The elucidation of the structure of the double helix provided a hint as to how DNA divides and makes copies of itself. This model suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied. What was not clear was how the replication took place. There were three models suggested ([\[link\]](#)): conservative, semi-conservative, and dispersive.

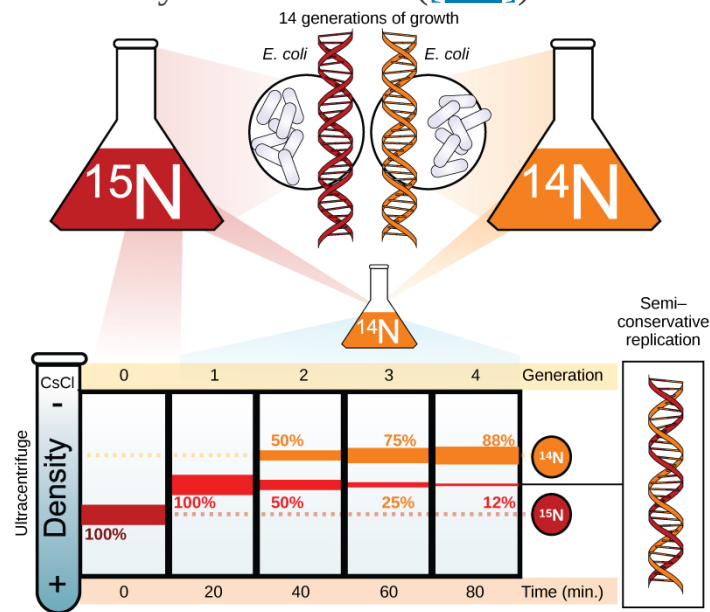


The three suggested models of DNA replication. Grey indicates the original DNA strands, and blue indicates newly synthesized DNA.

In conservative replication, the parental DNA remains together, and the newly formed daughter strands are together. The semi-conservative method

suggests that each of the two parental DNA strands act as a template for new DNA to be synthesized; after replication, each double-stranded DNA includes one parental or “old” strand and one “new” strand. In the dispersive model, both copies of DNA have double-stranded segments of parental DNA and newly synthesized DNA interspersed.

Meselson and Stahl were interested in understanding how DNA replicates. They grew *E. coli* for several generations in a medium containing a “heavy” isotope of nitrogen ( $^{15}\text{N}$ ) that gets incorporated into nitrogenous bases, and eventually into the DNA ([link](#)).



Meselson and Stahl experimented with *E. coli* grown first in heavy nitrogen ( $^{15}\text{N}$ ) then in  $^{14}\text{N}$ . DNA grown in  $^{15}\text{N}$  (red band) is heavier than DNA grown in  $^{14}\text{N}$  (orange band), and sediments to a lower level in cesium chloride solution in an ultracentrifuge. When DNA grown in  $^{15}\text{N}$  is switched to media containing  $^{14}\text{N}$ , after one round of cell division the DNA sediments halfway between the  $^{15}\text{N}$  and  $^{14}\text{N}$  levels, indicating that it now contains fifty percent  $^{14}\text{N}$ . In

subsequent cell divisions, an increasing amount of DNA contains  $^{14}\text{N}$  only. This data supports the semi-conservative replication model.  
(credit: modification of work by Mariana Ruiz Villareal)

The *E. coli* culture was then shifted into medium containing  $^{14}\text{N}$  and allowed to grow for one generation. The cells were harvested and the DNA was isolated. The DNA was centrifuged at high speeds in an ultracentrifuge. Some cells were allowed to grow for one more life cycle in  $^{14}\text{N}$  and spun again. During the density gradient centrifugation, the DNA is loaded into a gradient (typically a salt such as cesium chloride or sucrose) and spun at high speeds of 50,000 to 60,000 rpm. Under these circumstances, the DNA will form a band according to its density in the gradient. DNA grown in  $^{15}\text{N}$  will band at a higher density position than that grown in  $^{14}\text{N}$ . Meselson and Stahl noted that after one generation of growth in  $^{14}\text{N}$  after they had been shifted from  $^{15}\text{N}$ , the single band observed was intermediate in position in between DNA of cells grown exclusively in  $^{15}\text{N}$  and  $^{14}\text{N}$ . This suggested either a semi-conservative or dispersive mode of replication. The DNA harvested from cells grown for two generations in  $^{14}\text{N}$  formed two bands: one DNA band was at the intermediate position between  $^{15}\text{N}$  and  $^{14}\text{N}$ , and the other corresponded to the band of  $^{14}\text{N}$  DNA. These results could only be explained if DNA replicates in a semi-conservative manner. Therefore, the other two modes were ruled out.

During DNA replication, each of the two strands that make up the double helix serves as a template from which new strands are copied. The new strand will be complementary to the parental or “old” strand. When two daughter DNA copies are formed, they have the same sequence and are divided equally into the two daughter cells.

**Note:**

## Link to Learning



Click through [this tutorial](#) on DNA replication.

## Section Summary

The model for DNA replication suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied. In conservative replication, the parental DNA is conserved, and the daughter DNA is newly synthesized. The semi-conservative method suggests that each of the two parental DNA strands acts as template for new DNA to be synthesized; after replication, each double-stranded DNA includes one parental or “old” strand and one “new” strand. The dispersive mode suggested that the two copies of the DNA would have segments of parental DNA and newly synthesized DNA.

## Review Questions

### Exercise:

#### **Problem:**

Meselson and Stahl's experiments proved that DNA replicates by which mode?

- a. conservative
- b. semi-conservative
- c. dispersive
- d. none of the above

---

**Solution:**

B

**Exercise:****Problem:**

If the sequence of the 5'-3' strand is AATGCTAC, then the complementary sequence has the following sequence:

- a. 3'-AATGCTAC-5'
- b. 3'-CATCGTAA-5'
- c. 3'-TTACGATG-5'
- d. 3'-GTAGCATT-5'

---

**Solution:**

C

**Free Response****Exercise:****Problem:**

How did the scientific community learn that DNA replication takes place in a semi-conservative fashion?

---

**Solution:**

Meselson's experiments with *E. coli* grown in  $^{15}\text{N}$  deduced this finding.

## DNA Replication in Prokaryotes

By the end of this section, you will be able to:

- Explain the process of DNA replication in prokaryotes
- Discuss the role of different enzymes and proteins in supporting this process

DNA replication has been extremely well studied in prokaryotes primarily because of the small size of the genome and the mutants that are available. *E. coli* has 4.6 million base pairs in a single circular chromosome and all of it gets replicated in approximately 42 minutes, starting from a single origin of replication and proceeding around the circle in both directions. This means that approximately 1000 nucleotides are added per second. The process is quite rapid and occurs without many mistakes.

DNA replication employs a large number of proteins and enzymes, each of which plays a critical role during the process. One of the key players is the enzyme DNA polymerase, also known as DNA pol, which adds nucleotides one by one to the growing DNA chain that are complementary to the template strand. The addition of nucleotides requires energy; this energy is obtained from the nucleotides that have three phosphates attached to them, similar to ATP which has three phosphate groups attached. When the bond between the phosphates is broken, the energy released is used to form the phosphodiester bond between the incoming nucleotide and the growing chain. In prokaryotes, three main types of polymerases are known: DNA pol I, DNA pol II, and DNA pol III. It is now known that DNA pol III is the enzyme required for DNA synthesis; DNA pol I and DNA pol II are primarily required for repair.

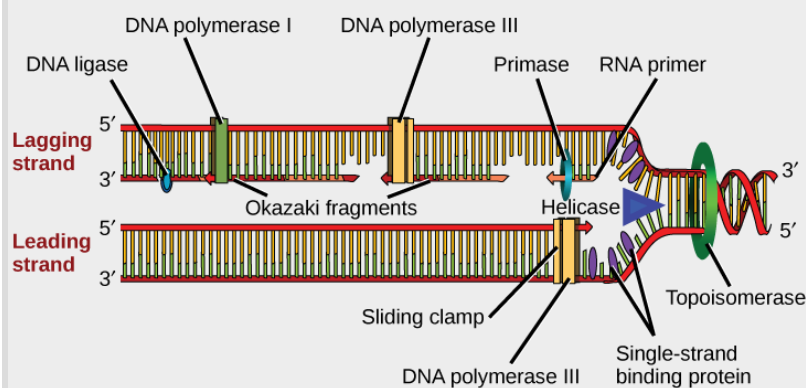
How does the replication machinery know where to begin? It turns out that there are specific nucleotide sequences called origins of replication where replication begins. In *E. coli*, which has a single origin of replication on its one chromosome (as do most prokaryotes), it is approximately 245 base pairs long and is rich in AT sequences. The origin of replication is recognized by certain proteins that bind to this site. An enzyme called **helicase** unwinds the DNA by breaking the hydrogen bonds between the nitrogenous base pairs. ATP hydrolysis is required for this process. As the DNA opens up, Y-shaped structures called **replication forks** are formed.



Two replication forks are formed at the origin of replication and these get extended bi- directionally as replication proceeds. **Single-strand binding proteins** coat the single strands of DNA near the replication fork to prevent the single-stranded DNA from winding back into a double helix. DNA polymerase is able to add nucleotides only in the 5' to 3' direction (a new DNA strand can be only extended in this direction). It also requires a free 3'-OH group to which it can add nucleotides by forming a phosphodiester bond between the 3'-OH end and the 5' phosphate of the next nucleotide. This essentially means that it cannot add nucleotides if a free 3'-OH group is not available. Then how does it add the first nucleotide? The problem is solved with the help of a primer that provides the free 3'-OH end. Another enzyme, RNA **primase**, synthesizes an RNA primer that is about five to ten nucleotides long and complementary to the DNA. Because this sequence primes the DNA synthesis, it is appropriately called the **primer**. DNA polymerase can now extend this RNA primer, adding nucleotides one by one that are complementary to the template strand ([\[link\]](#)).

### Note:

#### Art Connection



A replication fork is formed when helicase separates the DNA strands at the origin of replication. The DNA tends to become more highly coiled ahead of the replication fork. Topoisomerase breaks and reforms DNA's phosphate backbone ahead of the replication fork, thereby relieving the pressure that

results from this supercoiling. Single-strand binding proteins bind to the single-stranded DNA to prevent the helix from re-forming. Primase synthesizes an RNA primer. DNA polymerase III uses this primer to synthesize the daughter DNA strand. On the leading strand, DNA is synthesized continuously, whereas on the lagging strand, DNA is synthesized in short stretches called Okazaki fragments. DNA polymerase I replaces the RNA primer with DNA. DNA ligase seals the gaps between the Okazaki fragments, joining the fragments into a single DNA molecule. (credit: modification of work by Mariana Ruiz Villareal)

You isolate a cell strain in which the joining together of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

The replication fork moves at the rate of 1000 nucleotides per second. DNA polymerase can only extend in the 5' to 3' direction, which poses a slight problem at the replication fork. As we know, the DNA double helix is anti-parallel; that is, one strand is in the 5' to 3' direction and the other is oriented in the 3' to 5' direction. One strand, which is complementary to the 3' to 5' parental DNA strand, is synthesized continuously towards the replication fork because the polymerase can add nucleotides in this direction. This continuously synthesized strand is known as the **leading strand**. The other strand, complementary to the 5' to 3' parental DNA, is extended away from the replication fork, in small fragments known as **Okazaki fragments**, each requiring a primer to start the synthesis. Okazaki fragments are named after the Japanese scientist who first discovered them. The strand with the Okazaki fragments is known as the **lagging strand**.

The leading strand can be extended by one primer alone, whereas the lagging strand needs a new primer for each of the short Okazaki fragments. The overall direction of the lagging strand will be 3' to 5', and that of the leading strand 5' to 3'. A protein called the **sliding clamp** holds the DNA polymerase in place as it continues to add nucleotides. The sliding clamp is a ring-shaped protein that binds to the DNA and holds the polymerase in place. **Topoisomerase** prevents the over-winding of the DNA double helix ahead of the replication fork as the DNA is opening up; it does so by causing temporary nicks in the DNA helix and then resealing it. As synthesis proceeds, the RNA primers are replaced by DNA. The primers are removed by the exonuclease activity of DNA pol I, and the gaps are filled in by deoxyribonucleotides. The nicks that remain between the newly synthesized DNA (that replaced the RNA primer) and the previously synthesized DNA are sealed by the enzyme DNA **ligase** that catalyzes the formation of phosphodiester linkage between the 3'-OH end of one nucleotide and the 5' phosphate end of the other fragment.

Once the chromosome has been completely replicated, the two DNA copies move into two different cells during cell division. The process of DNA replication can be summarized as follows:

1. DNA unwinds at the origin of replication.
2. Helicase opens up the DNA-forming replication forks; these are extended bidirectionally.
3. Single-strand binding proteins coat the DNA around the replication fork to prevent rewinding of the DNA.
4. Topoisomerase binds at the region ahead of the replication fork to prevent supercoiling.
5. Primase synthesizes RNA primers complementary to the DNA strand.
6. DNA polymerase starts adding nucleotides to the 3'-OH end of the primer.
7. Elongation of both the lagging and the leading strand continues.
8. RNA primers are removed by exonuclease activity.
9. Gaps are filled by DNA pol by adding dNTPs.
10. The gap between the two DNA fragments is sealed by DNA ligase, which helps in the formation of phosphodiester bonds.

[\[link\]](#) summarizes the enzymes involved in prokaryotic DNA replication and the functions of each.

<b>Prokaryotic DNA Replication: Enzymes and Their Function</b>	
<b>Enzyme/protein</b>	<b>Specific Function</b>
DNA pol I	Exonuclease activity removes RNA primer and replaces with newly synthesized DNA
DNA pol II	Repair function
DNA pol III	Main enzyme that adds nucleotides in the 5'-3' direction
Helicase	Opens the DNA helix by breaking hydrogen bonds between the nitrogenous bases
Ligase	Seals the gaps between the Okazaki fragments to create one continuous DNA strand
Primase	Synthesizes RNA primers needed to start replication
Sliding Clamp	Helps to hold the DNA polymerase in place when nucleotides are being added
Topoisomerase	Helps relieve the stress on DNA when unwinding by causing breaks and then resealing the DNA

Prokaryotic DNA Replication: Enzymes and Their Function	
Enzyme/protein	Specific Function
Single-strand binding proteins (SSB)	Binds to single-stranded DNA to avoid DNA rewinding back.

### Note:

Link to Learning



Review the full process of DNA replication [here](#).

## Section Summary

Replication in prokaryotes starts from a sequence found on the chromosome called the origin of replication—the point at which the DNA opens up. Helicase opens up the DNA double helix, resulting in the formation of the replication fork. Single-strand binding proteins bind to the single-stranded DNA near the replication fork to keep the fork open. Primase synthesizes an RNA primer to initiate synthesis by DNA polymerase, which can add nucleotides only in the 5' to 3' direction. One strand is synthesized continuously in the direction of the replication fork; this is called the leading strand. The other strand is synthesized in a direction away from the replication fork, in short stretches of DNA known as Okazaki fragments. This strand is known as the lagging strand. Once replication is completed, the RNA primers are replaced by DNA nucleotides and the DNA is sealed.

with DNA ligase, which creates phosphodiester bonds between the 3'-OH of one end and the 5' phosphate of the other strand.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) You isolate a cell strain in which the joining together of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

---

#### Solution:

[\[link\]](#) DNA ligase, as this enzyme joins together Okazaki fragments.

## Review Questions

### Exercise:

#### Problem:

Which of the following components is not involved during the formation of the replication fork?

- a. single-strand binding proteins
  - b. helicase
  - c. origin of replication
  - d. ligase
- 

#### Solution:

D

### Exercise:

**Problem:** Which of the following does the enzyme primase synthesize?

- a. DNA primer
- b. RNA primer
- c. Okazaki fragments
- d. phosphodiester linkage

---

**Solution:**

B

**Exercise:**

**Problem:** In which direction does DNA replication take place?

- a. 5'-3'
- b. 3'-5'
- c. 5'
- d. 3'

---

**Solution:**

A

**Free Response**

**Exercise:**

**Problem:**

DNA replication is bidirectional and discontinuous; explain your understanding of those concepts.

---

**Solution:**

At an origin of replication, two replication forks are formed that are extended in two directions. On the lagging strand, Okazaki fragments are formed in a discontinuous manner.

**Exercise:**

**Problem:** What are Okazaki fragments and how they are formed?

---

**Solution:**

Short DNA fragments are formed on the lagging strand synthesized in a direction away from the replication fork. These are synthesized by DNA pol.

**Exercise:**

**Problem:**

If the rate of replication in a particular prokaryote is 900 nucleotides per second, how long would it take 1.2 million base pair genomes to make two copies?

---

**Solution:**

1333 seconds or 22.2 minutes.

**Exercise:**

**Problem:**

Explain the events taking place at the replication fork. If the gene for helicase is mutated, what part of replication will be affected?

---

**Solution:**

At the replication fork, the events taking place are helicase action, binding of single-strand binding proteins, primer synthesis, and synthesis of new strands. If there is a mutated helicase gene, the replication fork will not be extended.

**Exercise:**



**Problem:**

What is the role of a primer in DNA replication? What would happen if you forgot to add a primer in a tube containing the reaction mix for a DNA sequencing reaction?

---

**Solution:**

Primer provides a 3'-OH group for DNA pol to start adding nucleotides. There would be no reaction in the tube without a primer, and no bands would be visible on the electrophoresis.

**Glossary**

helicase

during replication, this enzyme helps to open up the DNA helix by breaking the hydrogen bonds

lagging strand

during replication, the strand that is replicated in short fragments and away from the replication fork

leading strand

strand that is synthesized continuously in the 5'-3' direction which is synthesized in the direction of the replication fork

ligase

enzyme that catalyzes the formation of a phosphodiester linkage between the 3' OH and 5' phosphate ends of the DNA

Okazaki fragment

DNA fragment that is synthesized in short stretches on the lagging strand

primase

enzyme that synthesizes the RNA primer; the primer is needed for DNA pol to start synthesis of a new DNA strand

primer

short stretch of nucleotides that is required to initiate replication; in the case of replication, the primer has RNA nucleotides

replication fork

Y-shaped structure formed during initiation of replication

single-strand binding protein

during replication, protein that binds to the single-stranded DNA; this helps in keeping the two strands of DNA apart so that they may serve as templates

sliding clamp

ring-shaped protein that holds the DNA pol on the DNA strand

topoisomerase

enzyme that causes underwinding or overwinding of DNA when DNA replication is taking place

## DNA Replication in Eukaryotes

By the end of this section, you will be able to:

- Discuss the similarities and differences between DNA replication in eukaryotes and prokaryotes
- State the role of telomerase in DNA replication

Eukaryotic genomes are much more complex and larger in size than prokaryotic genomes. The human genome has three billion base pairs per haploid set of chromosomes, and 6 billion base pairs are replicated during the S phase of the cell cycle. There are multiple origins of replication on the eukaryotic chromosome; humans can have up to 100,000 origins of replication. The rate of replication is approximately 100 nucleotides per second, much slower than prokaryotic replication. In yeast, which is a eukaryote, special sequences known as Autonomously Replicating Sequences (ARS) are found on the chromosomes. These are equivalent to the origin of replication in *E. coli*.

The number of DNA polymerases in eukaryotes is much more than prokaryotes: 14 are known, of which five are known to have major roles during replication and have been well studied. They are known as pol  $\alpha$ , pol  $\beta$ , pol  $\gamma$ , pol  $\delta$ , and pol  $\epsilon$ .

The essential steps of replication are the same as in prokaryotes. Before replication can start, the DNA has to be made available as template. Eukaryotic DNA is bound to basic proteins known as histones to form structures called nucleosomes. The chromatin (the complex between DNA and proteins) may undergo some chemical modifications, so that the DNA may be able to slide off the proteins or be accessible to the enzymes of the DNA replication machinery. At the origin of replication, a pre-replication complex is made with other initiator proteins. Other proteins are then recruited to start the replication process ([\[link\]](#)).

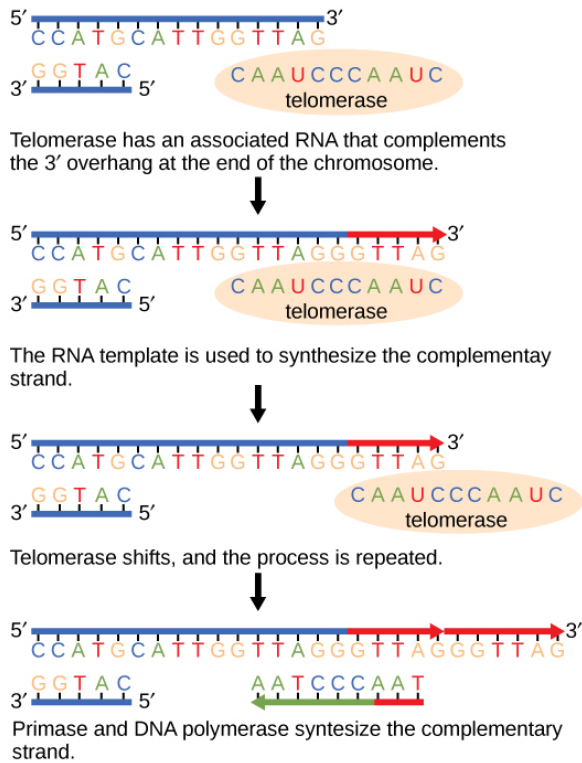
A helicase using the energy from ATP hydrolysis opens up the DNA helix. Replication forks are formed at each replication origin as the DNA unwinds. The opening of the double helix causes over-winding, or supercoiling, in the DNA ahead of the replication fork. These are resolved with the action of topoisomerases. Primers are formed by the enzyme

primase, and using the primer, DNA pol can start synthesis. While the leading strand is continuously synthesized by the enzyme pol  $\delta$ , the lagging strand is synthesized by pol  $\epsilon$ . A sliding clamp protein known as PCNA (Proliferating Cell Nuclear Antigen) holds the DNA pol in place so that it does not slide off the DNA. RNase H removes the RNA primer, which is then replaced with DNA nucleotides. The Okazaki fragments in the lagging strand are joined together after the replacement of the RNA primers with DNA. The gaps that remain are sealed by DNA ligase, which forms the phosphodiester bond.

## Telomere replication

Unlike prokaryotic chromosomes, eukaryotic chromosomes are linear. As you've learned, the enzyme DNA pol can add nucleotides only in the 5' to 3' direction. In the leading strand, synthesis continues until the end of the chromosome is reached. On the lagging strand, DNA is synthesized in short stretches, each of which is initiated by a separate primer. When the replication fork reaches the end of the linear chromosome, there is no place for a primer to be made for the DNA fragment to be copied at the end of the chromosome. These ends thus remain unpaired, and over time these ends may get progressively shorter as cells continue to divide.

The ends of the linear chromosomes are known as **telomeres**, which have repetitive sequences that code for no particular gene. In a way, these telomeres protect the genes from getting deleted as cells continue to divide. In humans, a six base pair sequence, TTAGGG, is repeated 100 to 1000 times. The discovery of the enzyme telomerase ([\[link\]](#)) helped in the understanding of how chromosome ends are maintained. The **telomerase** enzyme contains a catalytic part and a built-in RNA template. It attaches to the end of the chromosome, and complementary bases to the RNA template are added on the 3' end of the DNA strand. Once the 3' end of the lagging strand template is sufficiently elongated, DNA polymerase can add the nucleotides complementary to the ends of the chromosomes. Thus, the ends of the chromosomes are replicated.



The ends of linear chromosomes are maintained by the action of the telomerase enzyme.

Telomerase is typically active in germ cells and adult stem cells. It is not active in adult somatic cells. For her discovery of telomerase and its action, Elizabeth Blackburn ([\[link\]](#)) received the Nobel Prize for Medicine and Physiology in 2009.



Elizabeth Blackburn, 2009 Nobel Laureate, is the scientist who discovered how telomerase works.  
(credit: US Embassy Sweden)

## **Telomerase and Aging**

Cells that undergo cell division continue to have their telomeres shortened because most somatic cells do not make telomerase. This essentially means that telomere shortening is associated with aging. With the advent of modern medicine, preventative health care, and healthier lifestyles, the human life span has increased, and there is an increasing demand for people to look younger and have a better quality of life as they grow older.

In 2010, scientists found that telomerase can reverse some age-related conditions in mice. This may have potential in regenerative medicine. [\[footnote\]](#) Telomerase-deficient mice were used in these studies; these mice have tissue atrophy, stem cell depletion, organ system failure, and impaired tissue injury responses. Telomerase reactivation in these mice caused extension of telomeres, reduced DNA damage, reversed neurodegeneration, and improved the function of the testes, spleen, and intestines. Thus,

telomere reactivation may have potential for treating age-related diseases in humans.

Jaskelioff et al., “Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice,” *Nature* 469 (2011): 102-7.

Cancer is characterized by uncontrolled cell division of abnormal cells. The cells accumulate mutations, proliferate uncontrollably, and can migrate to different parts of the body through a process called metastasis. Scientists have observed that cancerous cells have considerably shortened telomeres and that telomerase is active in these cells. Interestingly, only after the telomeres were shortened in the cancer cells did the telomerase become active. If the action of telomerase in these cells can be inhibited by drugs during cancer therapy, then the cancerous cells could potentially be stopped from further division.

Difference between Prokaryotic and Eukaryotic Replication		
Property	Prokaryotes	Eukaryotes
Origin of replication	Single	Multiple
Rate of replication	1000 nucleotides/s	50 to 100 nucleotides/s
DNA polymerase types	5	14
Telomerase	Not present	Present
RNA primer removal	DNA pol I	RNase H
Strand elongation	DNA pol III	Pol $\delta$ , pol $\epsilon$

<b>Difference between Prokaryotic and Eukaryotic Replication</b>		
<b>Property</b>	<b>Prokaryotes</b>	<b>Eukaryotes</b>
Sliding clamp	Sliding clamp	PCNA

## Section Summary

Replication in eukaryotes starts at multiple origins of replication. The mechanism is quite similar to prokaryotes. A primer is required to initiate synthesis, which is then extended by DNA polymerase as it adds nucleotides one by one to the growing chain. The leading strand is synthesized continuously, whereas the lagging strand is synthesized in short stretches called Okazaki fragments. The RNA primers are replaced with DNA nucleotides; the DNA remains one continuous strand by linking the DNA fragments with DNA ligase. The ends of the chromosomes pose a problem as polymerase is unable to extend them without a primer. Telomerase, an enzyme with an inbuilt RNA template, extends the ends by copying the RNA template and extending one end of the chromosome. DNA polymerase can then extend the DNA using the primer. In this way, the ends of the chromosomes are protected.

## Review Questions

### Exercise:

**Problem:** The ends of the linear chromosomes are maintained by

- a. helicase
- b. primase
- c. DNA pol
- d. telomerase

---

**Solution:**



D

## Free Response

### Exercise:

#### Problem:

How do the linear chromosomes in eukaryotes ensure that its ends are replicated completely?

---

#### Solution:

Telomerase has an inbuilt RNA template that extends the 3' end, so primer is synthesized and extended. Thus, the ends are protected.

## Glossary

telomerase

enzyme that contains a catalytic part and an inbuilt RNA template; it functions to maintain telomeres at chromosome ends

telomere

DNA at the end of linear chromosomes

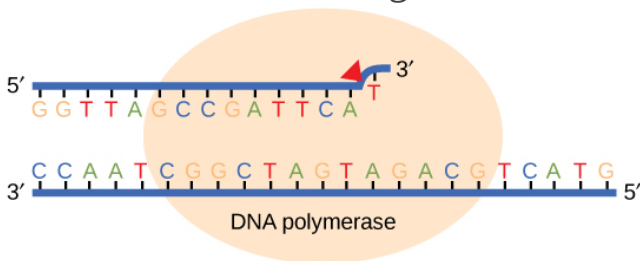
## DNA Repair

By the end of this section, you will be able to:

- Discuss the different types of mutations in DNA
- Explain DNA repair mechanisms

DNA replication is a highly accurate process, but mistakes can occasionally occur, such as a DNA polymerase inserting a wrong base. Uncorrected mistakes may sometimes lead to serious consequences, such as cancer. Repair mechanisms correct the mistakes. In rare cases, mistakes are not corrected, leading to mutations; in other cases, repair enzymes are themselves mutated or defective.

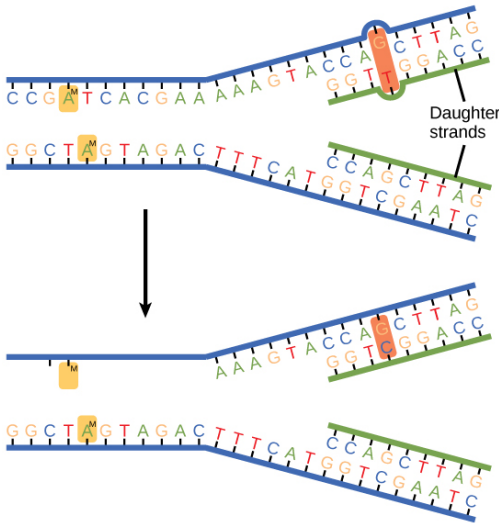
Most of the mistakes during DNA replication are promptly corrected by DNA polymerase by proofreading the base that has been just added ([\[link\]](#)). In **proofreading**, the DNA pol reads the newly added base before adding the next one, so a correction can be made. The polymerase checks whether the newly added base has paired correctly with the base in the template strand. If it is the right base, the next nucleotide is added. If an incorrect base has been added, the enzyme makes a cut at the phosphodiester bond and releases the wrong nucleotide. This is performed by the exonuclease action of DNA pol III. Once the incorrect nucleotide has been removed, a new one will be added again.



Proofreading by DNA polymerase corrects errors during replication.

Some errors are not corrected during replication, but are instead corrected after replication is completed; this type of repair is known as **mismatch repair** ([\[link\]](#)). The enzymes recognize the incorrectly added nucleotide and

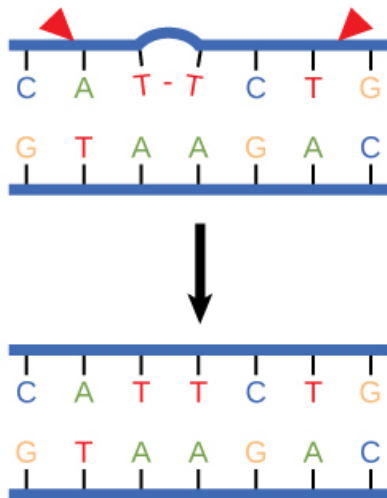
excise it; this is then replaced by the correct base. If this remains uncorrected, it may lead to more permanent damage. How do mismatch repair enzymes recognize which of the two bases is the incorrect one? In *E. coli*, after replication, the nitrogenous base adenine acquires a methyl group; the parental DNA strand will have methyl groups, whereas the newly synthesized strand lacks them. Thus, DNA polymerase is able to remove the wrongly incorporated bases from the newly synthesized, non-methylated strand. In eukaryotes, the mechanism is not very well understood, but it is believed to involve recognition of unsealed nicks in the new strand, as well as a short-term continuing association of some of the replication proteins with the new daughter strand after replication has completed.



In mismatch repair, the incorrectly added base is detected after replication.

The mismatch repair proteins detect this base and remove it from the newly synthesized strand by nuclease action. The gap is now filled with the correctly paired base.

In another type of repair mechanism, **nucleotide excision repair**, enzymes replace incorrect bases by making a cut on both the 3' and 5' ends of the incorrect base ([\[link\]](#)). The segment of DNA is removed and replaced with the correctly paired nucleotides by the action of DNA pol. Once the bases are filled in, the remaining gap is sealed with a phosphodiester linkage catalyzed by DNA ligase. This repair mechanism is often employed when UV exposure causes the formation of pyrimidine dimers.



Nucleotide excision repairs thymine dimers. When exposed to UV, thymine bases lying adjacent to each other can form thymine dimers. In normal cells, they are excised and replaced.

A well-studied example of mistakes not being corrected is seen in people suffering from xeroderma pigmentosa ([\[link\]](#)). Affected individuals have skin that is highly sensitive to UV rays from the sun. When individuals are exposed to UV, pyrimidine dimers, especially those of thymine, are formed;

people with xeroderma pigmentosa are not able to repair the damage. These are not repaired because of a defect in the nucleotide excision repair enzymes, whereas in normal individuals, the thymine dimers are excised and the defect is corrected. The thymine dimers distort the structure of the DNA double helix, and this may cause problems during DNA replication. People with xeroderma pigmentosa may have a higher risk of contracting skin cancer than those who don't have the condition.



Xeroderma pigmentosa is a condition in which thymine dimerization from exposure to UV is not repaired. Exposure to sunlight results in skin lesions.  
(credit: James Halpern et al.)

Errors during DNA replication are not the only reason why mutations arise in DNA. **Mutations**, variations in the nucleotide sequence of a genome, can also occur because of damage to DNA. Such mutations may be of two types: induced or spontaneous. **Induced mutations** are those that result from an exposure to chemicals, UV rays, x-rays, or some other environmental agent. **Spontaneous mutations** occur without any exposure to any environmental agent; they are a result of natural reactions taking place within the body.

Mutations may have a wide range of effects. Some mutations are not expressed; these are known as **silent mutations**. **Point mutations** are those mutations that affect a single base pair. The most common nucleotide mutations are substitutions, in which one base is replaced by another. These can be of two types, either transitions or transversions. **Transition substitution** refers to a purine or pyrimidine being replaced by a base of the same kind; for example, a purine such as adenine may be replaced by the purine guanine. **Transversion substitution** refers to a purine being replaced by a pyrimidine, or vice versa; for example, cytosine, a pyrimidine, is replaced by adenine, a purine. Mutations can also be the result of the addition of a base, known as an insertion, or the removal of a base, also known as deletion. Sometimes a piece of DNA from one chromosome may get translocated to another chromosome or to another region of the same chromosome; this is also known as translocation. These mutation types are shown in [\[link\]](#).

**Note:**

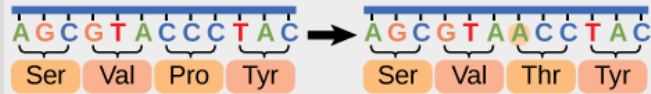
Art Connection

### Point Mutations

Silent: has no effect on the protein sequence



Missense: results in an amino acid substitution

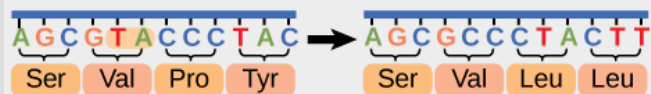


Nonsense: substitutes a stop codon for an amino acid



### Frameshift Mutations

Insertions or deletions of nucleotides may result in a shift in the reading frame or insertion of a stop codon.



Mutations can lead to changes in the protein sequence encoded by the DNA.

A frameshift mutation that results in the insertion of three nucleotides is often less deleterious than a mutation that results in the insertion of one nucleotide. Why?

Mutations in repair genes have been known to cause cancer. Many mutated repair genes have been implicated in certain forms of pancreatic cancer, colon cancer, and colorectal cancer. Mutations can affect either somatic cells or germ cells. If many mutations accumulate in a somatic cell, they may lead to problems such as the uncontrolled cell division observed in cancer. If a mutation takes place in germ cells, the mutation will be passed

on to the next generation, as in the case of hemophilia and xeroderma pigmentosa.

## Section Summary

DNA polymerase can make mistakes while adding nucleotides. It edits the DNA by proofreading every newly added base. Incorrect bases are removed and replaced by the correct base, and then a new base is added. Most mistakes are corrected during replication, although when this does not happen, the mismatch repair mechanism is employed. Mismatch repair enzymes recognize the wrongly incorporated base and excise it from the DNA, replacing it with the correct base. In yet another type of repair, nucleotide excision repair, the incorrect base is removed along with a few bases on the 5' and 3' end, and these are replaced by copying the template with the help of DNA polymerase. The ends of the newly synthesized fragment are attached to the rest of the DNA using DNA ligase, which creates a phosphodiester bond.

Most mistakes are corrected, and if they are not, they may result in a mutation defined as a permanent change in the DNA sequence. Mutations can be of many types, such as substitution, deletion, insertion, and translocation. Mutations in repair genes may lead to serious consequences such as cancer. Mutations can be induced or may occur spontaneously.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) A frameshift mutation that results in the insertion of three nucleotides is often less deleterious than a mutation that results in the insertion of one nucleotide. Why?

---

#### Solution:



[\[link\]](#) If three nucleotides are added, one additional amino acid will be incorporated into the protein chain, but the reading frame won't shift.

## Review Questions

### Exercise:

#### Problem:

During proofreading, which of the following enzymes reads the DNA?

- a. primase
- b. topoisomerase
- c. DNA pol
- d. helicase

---

#### Solution:

C

### Exercise:

#### Problem:

The initial mechanism for repairing nucleotide errors in DNA is \_\_\_\_\_.

- a. mismatch repair
- b. DNA polymerase proofreading
- c. nucleotide excision repair
- d. thymine dimers

---

#### Solution:

B

## Free Response

### Exercise:

#### Problem:

What is the consequence of mutation of a mismatch repair enzyme?  
How will this affect the function of a gene?

---

#### Solution:

Mutations are not repaired, as in the case of xeroderma pigmentosa.  
Gene function may be affected or it may not be expressed.

## Glossary

induced mutation

mutation that results from exposure to chemicals or environmental agents

mutation

variation in the nucleotide sequence of a genome

mismatch repair

type of repair mechanism in which mismatched bases are removed after replication

nucleotide excision repair

type of DNA repair mechanism in which the wrong base, along with a few nucleotides upstream or downstream, are removed

proofreading

function of DNA pol in which it reads the newly added base before adding the next one

point mutation

mutation that affects a single base

silent mutation

mutation that is not expressed

spontaneous mutation

mutation that takes place in the cells as a result of chemical reactions taking place naturally without exposure to any external agent

transition substitution

when a purine is replaced with a purine or a pyrimidine is replaced with another pyrimidine

transversion substitution

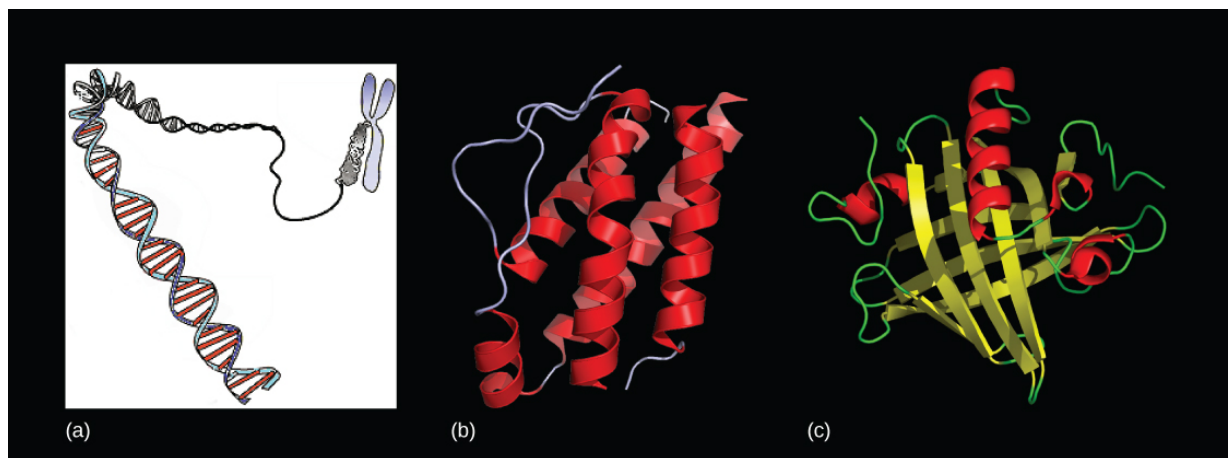
when a purine is replaced by a pyrimidine or a pyrimidine is replaced by a purine

## Introduction

class="introduction"

Genes, which are carried on (a) chromosomes, are linearly organized instructions for making the RNA and protein molecules that are necessary for all of processes of life. The (b) interleukin-2 protein and (c) alpha-2u-globulin protein are just two examples of the array of different molecular structures that are encoded by genes. (credit "chromosome: National Human Genome Research Institute; credit "interleukin-

2”: Ramin  
Herati/Created  
from PDB  
1M47 and  
rendered with  
Pymol; credit  
“alpha-2u-  
globulin”:  
Darren  
Logan/rende  
d with  
AISMIG)



Since the rediscovery of Mendel’s work in 1900, the definition of the gene has progressed from an abstract unit of heredity to a tangible molecular entity capable of replication, expression, and mutation ([\[link\]](#)). Genes are composed of DNA and are linearly arranged on chromosomes. Genes specify the sequences of amino acids, which are the building blocks of proteins. In turn, proteins are responsible for orchestrating nearly every function of the cell. Both genes and the proteins they encode are absolutely essential to life as we know it.

## The Genetic Code

By the end of this section, you will be able to:

- Explain the “central dogma” of protein synthesis
- Describe the genetic code and how the nucleotide sequence prescribes the amino acid and the protein sequence

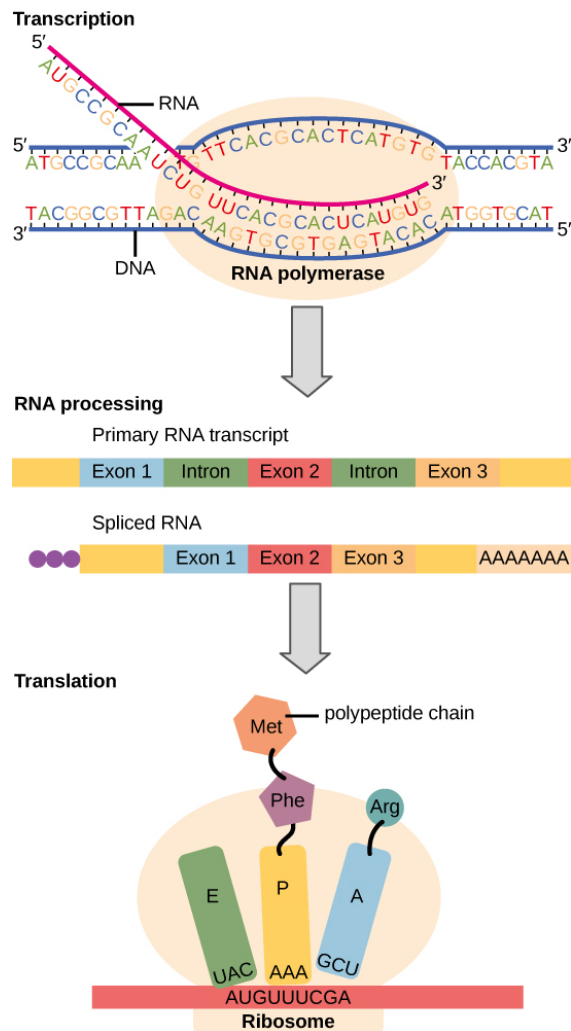
The cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template converts nucleotide-based genetic information into a protein product. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 letters ([link](#)). Each amino acid is defined by a three-nucleotide sequence called the triplet codon. Different amino acids have different chemistries (such as acidic versus basic, or polar and nonpolar) and different structural constraints. Variation in amino acid sequence gives rise to enormous variation in protein structure and function.

AMINO ACID			
Nonpolar, aliphatic R groups			
	Glycine	Alanine	Valine
	Leucine	Methionine	Isoleucine
	Serine	Threonine	Cysteine
Polar, uncharged R groups			
	Proline	Asparagine	Glutamine
AMINO ACID			
Positively charged R groups			
	Lysine	Arginine	Histidine
Negatively charged R groups			
	Aspartate	Glutamate	
Nonpolar, aromatic R groups			
	Phenylalanine	Tyrosine	Tryptophan

Structures of the 20 amino acids found in proteins are shown. Each amino acid is composed of an amino group ( $\text{NH}_3^+$ ), a carboxyl group ( $\text{COO}^-$ ), and a side chain (blue). The side chain may be nonpolar, polar, or charged, as well as large or small. It is the variety of amino acid side chains that gives rise to the incredible variation of protein structure and function.

## **The Central Dogma: DNA Encodes RNA; RNA Encodes Protein**

The flow of genetic information in cells from DNA to mRNA to protein is described by the **Central Dogma** ([\[link\]](#)), which states that genes specify the sequence of mRNAs, which in turn specify the sequence of proteins. The decoding of one molecule to another is performed by specific proteins and RNAs. Because the information stored in DNA is so central to cellular function, it makes intuitive sense that the cell would make mRNA copies of this information for protein synthesis, while keeping the DNA itself intact and protected. The copying of DNA to RNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every nucleotide read in the DNA strand. The translation to protein is a bit more complex because three mRNA nucleotides correspond to one amino acid in the polypeptide sequence. However, the translation to protein is still systematic and **colinear**, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.



Instructions on DNA are transcribed onto messenger RNA. Ribosomes are able to read the genetic information inscribed on a strand of messenger RNA and use this information to string amino acids together into a protein.

**The Genetic Code Is Degenerate and Universal**



Given the different numbers of “letters” in the mRNA and protein “alphabets,” scientists theorized that combinations of nucleotides corresponded to single amino acids. Nucleotide doublets would not be sufficient to specify every amino acid because there are only 16 possible two-nucleotide combinations ( $4^2$ ). In contrast, there are 64 possible nucleotide triplets ( $4^3$ ), which is far more than the number of amino acids. Scientists theorized that amino acids were encoded by nucleotide triplets and that the genetic code was **degenerate**. In other words, a given amino acid could be encoded by more than one nucleotide triplet. This was later confirmed experimentally; Francis Crick and Sydney Brenner used the chemical mutagen proflavin to insert one, two, or three nucleotides into the gene of a virus. When one or two nucleotides were inserted, protein synthesis was completely abolished. When three nucleotides were inserted, the protein was synthesized and functional. This demonstrated that three nucleotides specify each amino acid. These nucleotide triplets are called **codons**. The insertion of one or two nucleotides completely changed the triplet **reading frame**, thereby altering the message for every subsequent amino acid ([\[link\]](#)). Though insertion of three nucleotides caused an extra amino acid to be inserted during translation, the integrity of the rest of the protein was maintained.

Scientists painstakingly solved the genetic code by translating synthetic mRNAs in vitro and sequencing the proteins they specified ([\[link\]](#)).

		Second letter				
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G

This figure shows the genetic code for translating each nucleotide triplet in mRNA into an amino acid or a termination signal in a nascent protein.  
(credit: modification of work by NIH)

In addition to instructing the addition of a specific amino acid to a polypeptide chain, three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called **nonsense codons**, or stop codons. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the start codon to initiate translation. The reading frame for translation is set by the AUG start codon near the 5' end of the mRNA.

The genetic code is universal. With a few exceptions, virtually all species use the same genetic code for protein synthesis. Conservation of codons means that a purified mRNA encoding the globin protein in horses could be transferred to a tulip cell, and the tulip would synthesize horse globin. That there is only one genetic code is powerful evidence that all of life on Earth shares a common origin, especially considering that there are about  $10^{84}$  possible combinations of 20 amino acids and 64 triplet codons.

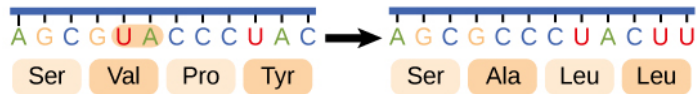
**Note:**

Link to Learning



Transcribe a gene and translate it to protein using complementary pairing and the genetic code at this [site](#).

#### Frameshift Mutations



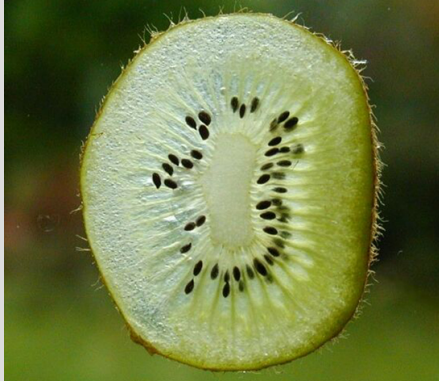
The deletion of two nucleotides shifts the reading frame of an mRNA and changes the entire protein message, creating a nonfunctional protein or terminating protein synthesis altogether.

Degeneracy is believed to be a cellular mechanism to reduce the negative impact of random mutations. Codons that specify the same amino acid typically only differ by one nucleotide. In addition, amino acids with chemically similar side chains are encoded by similar codons. This nuance of the genetic code ensures that a single-nucleotide substitution mutation might either specify the same amino acid but have no effect or specify a similar amino acid, preventing the protein from being rendered completely nonfunctional.

#### **Note:**

Scientific Method Connection

**Which Has More DNA: A Kiwi or a Strawberry?**



Do you think that a kiwi or a strawberry has more DNA per fruit? (credit “kiwi”: "Kelbv"/Flickr; credit: “strawberry”: Alisdair McDiarmid)

**Question:** Would a kiwifruit and strawberry that are approximately the same size ([link](#)) also have approximately the same amount of DNA?

**Background:** Genes are carried on chromosomes and are made of DNA. All mammals are diploid, meaning they have two copies of each chromosome. However, not all plants are diploid. The common strawberry is octoploid ( $8n$ ) and the cultivated kiwi is hexaploid ( $6n$ ). Research the total number of chromosomes in the cells of each of these fruits and think about how this might correspond to the amount of DNA in these fruits' cell nuclei. Read about the technique of DNA isolation to understand how each step in the isolation protocol helps liberate and precipitate DNA.

**Hypothesis:** Hypothesize whether you would be able to detect a difference in DNA quantity from similarly sized strawberries and kiwis. Which fruit do you think would yield more DNA?

**Test your hypothesis:** Isolate the DNA from a strawberry and a kiwi that are similarly sized. Perform the experiment in at least triplicate for each fruit.

1. Prepare a bottle of DNA extraction buffer from 900 mL water, 50 mL dish detergent, and two teaspoons of table salt. Mix by inversion (cap it and turn it upside down a few times).
2. Grind a strawberry and a kiwifruit by hand in a plastic bag, or using a mortar and pestle, or with a metal bowl and the end of a blunt instrument. Grind for at least two minutes per fruit.

3. Add 10 mL of the DNA extraction buffer to each fruit, and mix well for at least one minute.
4. Remove cellular debris by filtering each fruit mixture through cheesecloth or porous cloth and into a funnel placed in a test tube or an appropriate container.
5. Pour ice-cold ethanol or isopropanol (rubbing alcohol) into the test tube. You should observe white, precipitated DNA.
6. Gather the DNA from each fruit by winding it around separate glass rods.

**Record your observations:** Because you are not quantitatively measuring DNA volume, you can record for each trial whether the two fruits produced the same or different amounts of DNA as observed by eye. If one or the other fruit produced noticeably more DNA, record this as well. Determine whether your observations are consistent with several pieces of each fruit.

**Analyze your data:** Did you notice an obvious difference in the amount of DNA produced by each fruit? Were your results reproducible?

**Draw a conclusion:** Given what you know about the number of chromosomes in each fruit, can you conclude that chromosome number necessarily correlates to DNA amount? Can you identify any drawbacks to this procedure? If you had access to a laboratory, how could you standardize your comparison and make it more quantitative?

## Section Summary

The genetic code refers to the DNA alphabet (A, T, C, G), the RNA alphabet (A, U, C, G), and the polypeptide alphabet (20 amino acids). The Central Dogma describes the flow of genetic information in the cell from genes to mRNA to proteins. Genes are used to make mRNA by the process of transcription; mRNA is used to synthesize proteins by the process of translation. The genetic code is degenerate because 64 triplet codons in mRNA specify only 20 amino acids and three nonsense codons. Almost every species on the planet uses the same genetic code.

## Review Questions

### Exercise:

#### Problem:

The AUC and AUA codons in mRNA both specify isoleucine. What feature of the genetic code explains this?

- a. complementarity
- b. nonsense codons
- c. universality
- d. degeneracy

---

#### Solution:

D

### Exercise:

**Problem:**How many nucleotides are in 12 mRNA codons?

- a. 12
- b. 24
- c. 36
- d. 48

---

#### Solution:

C

## Free Response

### Exercise:

**Problem:**

Imagine if there were 200 commonly occurring amino acids instead of 20. Given what you know about the genetic code, what would be the shortest possible codon length? Explain.

---

**Solution:**

For 200 commonly occurring amino acids, codons consisting of four types of nucleotides would have to be at least four nucleotides long, because  $4^4 = 256$ . There would be much less degeneracy in this case.

**Exercise:****Problem:**

Discuss how degeneracy of the genetic code makes cells more robust to mutations.

---

**Solution:**

Codons that specify the same amino acid typically only differ by one nucleotide. In addition, amino acids with chemically similar side chains are encoded by similar codons. This nuance of the genetic code ensures that a single-nucleotide substitution mutation might either specify the same amino acid and have no effect, or may specify a similar amino acid, preventing the protein from being rendered completely nonfunctional.

**Glossary****Central Dogma**

states that genes specify the sequence of mRNAs, which in turn specify the sequence of proteins

**codon**

three consecutive nucleotides in mRNA that specify the insertion of an amino acid or the release of a polypeptide chain during translation

colinear

in terms of RNA and protein, three “units” of RNA (nucleotides) specify one “unit” of protein (amino acid) in a consecutive fashion

degeneracy

(of the genetic code) describes that a given amino acid can be encoded by more than one nucleotide triplet; the code is degenerate, but not ambiguous

nonsense codon

one of the three mRNA codons that specifies termination of translation

reading frame

sequence of triplet codons in mRNA that specify a particular protein; a ribosome shift of one or two nucleotides in either direction completely abolishes synthesis of that protein



## Prokaryotic Transcription

By the end of this section, you will be able to:

- List the different steps in prokaryotic transcription
- Discuss the role of promoters in prokaryotic transcription
- Describe how and when transcription is terminated

The prokaryotes, which include bacteria and archaea, are mostly single-celled organisms that, by definition, lack membrane-bound nuclei and other organelles. A bacterial chromosome is a covalently closed circle that, unlike eukaryotic chromosomes, is not organized around histone proteins. The central region of the cell in which prokaryotic DNA resides is called the nucleoid. In addition, prokaryotes often have abundant **plasmids**, which are shorter circular DNA molecules that may only contain one or a few genes. Plasmids can be transferred independently of the bacterial chromosome during cell division and often carry traits such as antibiotic resistance.

Transcription in prokaryotes (and in eukaryotes) requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a **transcription bubble**. Transcription always proceeds from the same DNA strand for each gene, which is called the **template strand**. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the **nontemplate strand**. The only difference is that in mRNA, all of the T nucleotides are replaced with U nucleotides. In an RNA double helix, A can bind U via two hydrogen bonds, just as in A–T pairing in a DNA double helix.

The nucleotide pair in the DNA double helix that corresponds to the site from which the first 5' mRNA nucleotide is transcribed is called the +1 site, or the **initiation site**. Nucleotides preceding the initiation site are given negative numbers and are designated **upstream**. Conversely, nucleotides following the initiation site are denoted with “+” numbering and are called **downstream** nucleotides.

## Initiation of Transcription in Prokaryotes

Prokaryotes do not have membrane-enclosed nuclei. Therefore, the processes of transcription, translation, and mRNA degradation can all occur simultaneously. The intracellular level of a bacterial protein can quickly be amplified by multiple transcription and translation events occurring concurrently on the same DNA template. Prokaryotic transcription often covers more than one gene and produces polycistronic mRNAs that specify more than one protein.

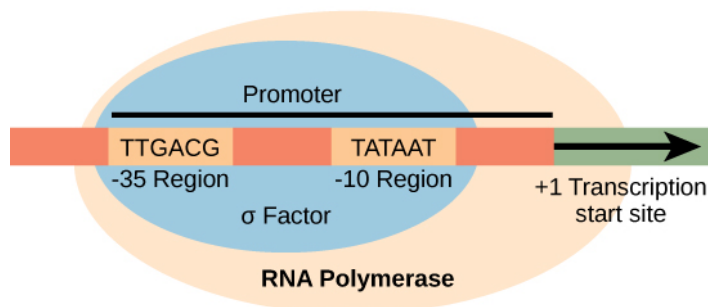
Our discussion here will exemplify transcription by describing this process in *Escherichia coli*, a well-studied bacterial species. Although some differences exist between transcription in *E. coli* and transcription in archaea, an understanding of *E. coli* transcription can be applied to virtually all bacterial species.

## Prokaryotic RNA Polymerase

Prokaryotes use the same RNA polymerase to transcribe all of their genes. In *E. coli*, the polymerase is composed of five polypeptide subunits, two of which are identical. Four of these subunits, denoted  $\alpha$ ,  $\alpha$ ,  $\beta$ , and  $\beta'$  comprise the polymerase **core enzyme**. These subunits assemble every time a gene is transcribed, and they disassemble once transcription is complete. Each subunit has a unique role; the two  $\alpha$ -subunits are necessary to assemble the polymerase on the DNA; the  $\beta$ -subunit binds to the ribonucleoside triphosphate that will become part of the nascent “recently born” mRNA molecule; and the  $\beta'$  binds the DNA template strand. The fifth subunit,  $\sigma$ , is involved only in transcription initiation. It confers transcriptional specificity such that the polymerase begins to synthesize mRNA from an appropriate initiation site. Without  $\sigma$ , the core enzyme would transcribe from random sites and would produce mRNA molecules that specified protein gibberish. The polymerase comprised of all five subunits is called the **holoenzyme**.

## Prokaryotic Promoters

A **promoter** is a DNA sequence onto which the transcription machinery binds and initiates transcription. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all the time, some of the time, or infrequently. Although promoters vary among prokaryotic genomes, a few elements are conserved. At the -10 and -35 regions upstream of the initiation site, there are two promoter **consensus** sequences, or regions that are similar across all promoters and across various bacterial species ([\[link\]](#)). The -10 consensus sequence, called the -10 region, is TATAAT. The -35 sequence, TTGACA, is recognized and bound by  $\sigma$ . Once this interaction is made, the subunits of the core enzyme bind to the site. The A–T-rich -10 region facilitates unwinding of the DNA template, and several phosphodiester bonds are made. The transcription initiation phase ends with the production of abortive transcripts, which are polymers of approximately 10 nucleotides that are made and released.



The  $\sigma$  subunit of prokaryotic RNA polymerase recognizes consensus sequences found in the promoter region upstream of the transcription start sight. The  $\sigma$  subunit dissociates from the polymerase after transcription has been initiated.

**Note:**

## Link to Learning

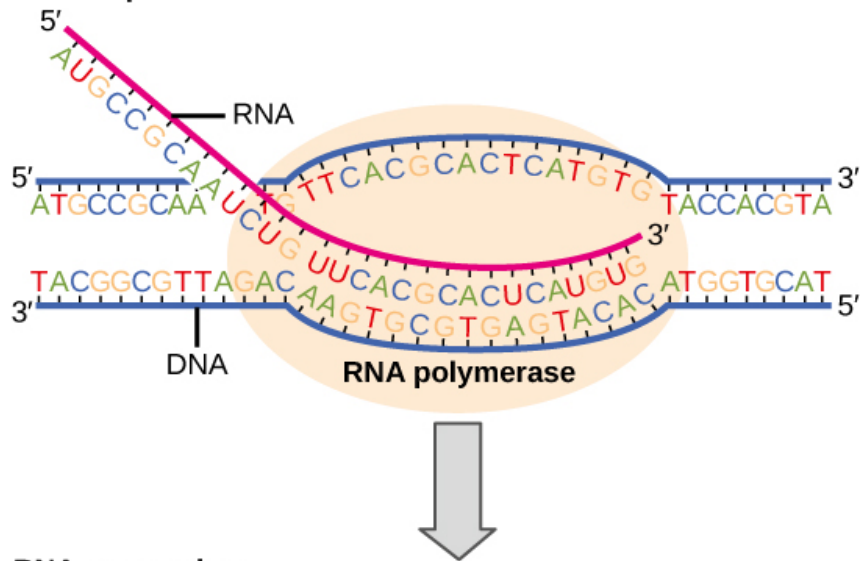


View this [MolecularMovies animation](#) to see the first part of transcription and the base sequence repetition of the TATA box.

## Elongation and Termination in Prokaryotes

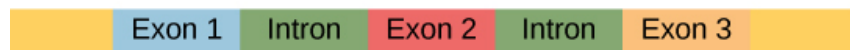
The transcription elongation phase begins with the release of the  $\sigma$  subunit from the polymerase. The dissociation of  $\sigma$  allows the core enzyme to proceed along the DNA template, synthesizing mRNA in the 5' to 3' direction at a rate of approximately 40 nucleotides per second. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it ([\[link\]](#)). The base pairing between DNA and RNA is not stable enough to maintain the stability of the mRNA synthesis components. Instead, the RNA polymerase acts as a stable linker between the DNA template and the nascent RNA strands to ensure that elongation is not interrupted prematurely.

## Transcription

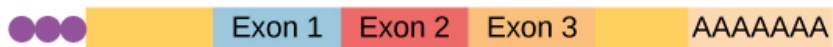


## RNA processing

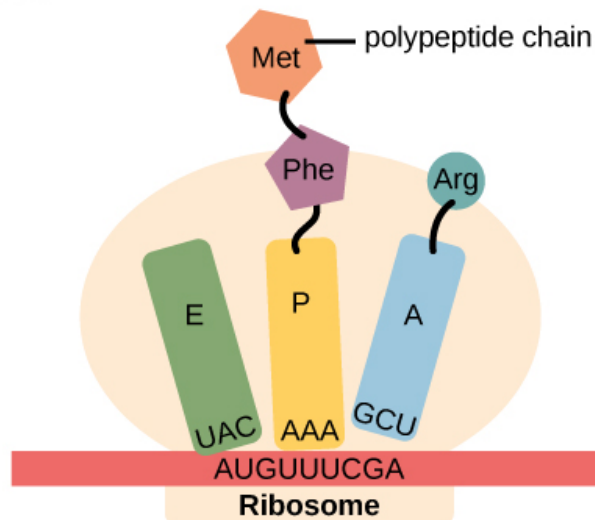
Primary RNA transcript



Spliced RNA



## Translation



During elongation, the prokaryotic RNA polymerase tracks along the DNA template,

synthesizes mRNA in the 5' to 3' direction, and unwinds and rewinds the DNA as it is read.

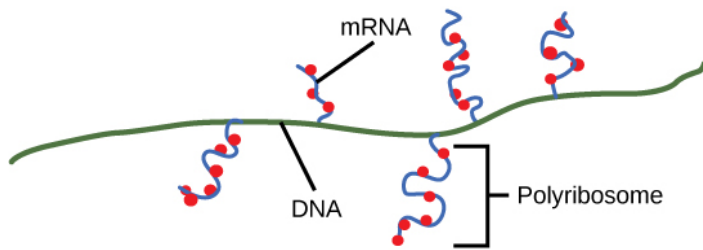
## Prokaryotic Termination Signals

Once a gene is transcribed, the prokaryotic polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals. One is protein-based and the other is RNA-based. **Rho-dependent termination** is controlled by the rho protein, which tracks along behind the polymerase on the growing mRNA chain. Near the end of the gene, the polymerase encounters a run of G nucleotides on the DNA template and it stalls. As a result, the rho protein collides with the polymerase. The interaction with rho releases the mRNA from the transcription bubble.

**Rho-independent termination** is controlled by specific sequences in the DNA template strand. As the polymerase nears the end of the gene being transcribed, it encounters a region rich in C–G nucleotides. The mRNA folds back on itself, and the complementary C–G nucleotides bind together. The result is a stable **hairpin** that causes the polymerase to stall as soon as it begins to transcribe a region rich in A–T nucleotides. The complementary U–A region of the mRNA transcript forms only a weak interaction with the template DNA. This, coupled with the stalled polymerase, induces enough instability for the core enzyme to break away and liberate the new mRNA transcript.

Upon termination, the process of transcription is complete. By the time termination occurs, the prokaryotic transcript would already have been used to begin synthesis of numerous copies of the encoded protein because these processes can occur concurrently. The unification of transcription, translation, and even mRNA degradation is possible because all of these processes occur in the same 5' to 3' direction, and because there is no membranous compartmentalization in the prokaryotic cell ([\[link\]](#)). In

contrast, the presence of a nucleus in eukaryotic cells precludes simultaneous transcription and translation.



Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides.

In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

**Note:**

Link to Learning



Visit this [BioStudio animation](#) to see the process of prokaryotic transcription.

## Section Summary

In prokaryotes, mRNA synthesis is initiated at a promoter sequence on the DNA template comprising two consensus sequences that recruit RNA polymerase. The prokaryotic polymerase consists of a core enzyme of four protein subunits and a  $\sigma$  protein that assists only with initiation. Elongation synthesizes mRNA in the 5' to 3' direction at a rate of 40 nucleotides per second. Termination liberates the mRNA and occurs either by rho protein interaction or by the formation of an mRNA hairpin.

## Review Questions

### Exercise:

#### Problem:

Which subunit of the *E. coli* polymerase confers specificity to transcription?

- a.  $\alpha$
- b.  $\beta$
- c.  $\beta'$
- d.  $\sigma$

---

#### Solution:

D

### Exercise:

#### Problem:

The -10 and -35 regions of prokaryotic promoters are called consensus sequences because \_\_\_\_\_.

- a. they are identical in all bacterial species
- b. they are similar in all bacterial species
- c. they exist in all organisms



d. they have the same function in all organisms

---

**Solution:**

B

**Free Response**

**Exercise:**

**Problem:**

If mRNA is complementary to the DNA template strand and the DNA template strand is complementary to the DNA nontemplate strand, then why are base sequences of mRNA and the DNA nontemplate strand not identical? Could they ever be?

---

**Solution:**

DNA is different from RNA in that T nucleotides in DNA are replaced with U nucleotides in RNA. Therefore, they could never be identical in base sequence.

**Exercise:**

**Problem:**

In your own words, describe the difference between rho-dependent and rho-independent termination of transcription in prokaryotes.

---

**Solution:**

Rho-dependent termination is controlled by the rho protein, which tracks along behind the polymerase on the growing mRNA chain. Near the end of the gene, the polymerase stalls at a run of G nucleotides on the DNA template. The rho protein collides with the polymerase and releases mRNA from the transcription bubble. Rho-independent termination is controlled by specific sequences in the DNA template

strand. As the polymerase nears the end of the gene being transcribed, it encounters a region rich in C–G nucleotides. This creates an mRNA hairpin that causes the polymerase to stall right as it begins to transcribe a region rich in A–T nucleotides. Because A–U bonds are less thermostable, the core enzyme falls away.

## Glossary

### consensus

DNA sequence that is used by many species to perform the same or similar functions

### core enzyme

prokaryotic RNA polymerase consisting of  $\alpha$ ,  $\alpha$ ,  $\beta$ , and  $\beta'$  but missing  $\sigma$ ; this complex performs elongation

### downstream

nucleotides following the initiation site in the direction of mRNA transcription; in general, sequences that are toward the 3' end relative to a site on the mRNA

### hairpin

structure of RNA when it folds back on itself and forms intramolecular hydrogen bonds between complementary nucleotides

### holoenzyme

prokaryotic RNA polymerase consisting of  $\alpha$ ,  $\alpha$ ,  $\beta$ ,  $\beta'$ , and  $\sigma$ ; this complex is responsible for transcription initiation

### initiation site

nucleotide from which mRNA synthesis proceeds in the 5' to 3' direction; denoted with a “+1”

### nontemplate strand

strand of DNA that is not used to transcribe mRNA; this strand is identical to the mRNA except that T nucleotides in the DNA are replaced by U nucleotides in the mRNA

plasmid

extrachromosomal, covalently closed, circular DNA molecule that may only contain one or a few genes; common in prokaryotes

promoter

DNA sequence to which RNA polymerase and associated factors bind and initiate transcription

Rho-dependent termination

in prokaryotes, termination of transcription by an interaction between RNA polymerase and the rho protein at a run of G nucleotides on the DNA template

Rho-independent

termination sequence-dependent termination of prokaryotic mRNA synthesis; caused by hairpin formation in the mRNA that stalls the polymerase

TATA box

conserved promoter sequence in eukaryotes and prokaryotes that helps to establish the initiation site for transcription

template strand

strand of DNA that specifies the complementary mRNA molecule

transcription bubble

region of locally unwound DNA that allows for transcription of mRNA

upstream

nucleotides preceding the initiation site; in general, sequences toward the 5' end relative to a site on the mRNA

## Eukaryotic Transcription

By the end of this section, you will be able to:

- List the steps in eukaryotic transcription
- Discuss the role of RNA polymerases in transcription
- Compare and contrast the three RNA polymerases
- Explain the significance of transcription factors

Prokaryotes and eukaryotes perform fundamentally the same process of transcription, with a few key differences. The most important difference between prokaryotes and eukaryotes is the latter's membrane-bound nucleus and organelles. With the genes bound in a nucleus, the eukaryotic cell must be able to transport its mRNA to the cytoplasm and must protect its mRNA from degrading before it is translated. Eukaryotes also employ three different polymerases that each transcribe a different subset of genes. Eukaryotic mRNAs are usually monogenic, meaning that they specify a single protein.

### Initiation of Transcription in Eukaryotes

Unlike the prokaryotic polymerase that can bind to a DNA template on its own, eukaryotes require several other proteins, called transcription factors, to first bind to the promoter region and then help recruit the appropriate polymerase.

### The Three Eukaryotic RNA Polymerases

The features of eukaryotic mRNA synthesis are markedly more complex those of prokaryotes. Instead of a single polymerase comprising five subunits, the eukaryotes have three polymerases that are each made up of 10 subunits or more. Each eukaryotic polymerase also requires a distinct set of transcription factors to bring it to the DNA template.

RNA polymerase I is located in the nucleolus, a specialized nuclear substructure in which ribosomal RNA (rRNA) is transcribed, processed, and assembled into ribosomes ([\[link\]](#)). The rRNA molecules are considered

structural RNAs because they have a cellular role but are not translated into protein. The rRNAs are components of the ribosome and are essential to the process of translation. RNA polymerase I synthesizes all of the rRNAs except for the 5S rRNA molecule. The “S” designation applies to “Svedberg” units, a nonadditive value that characterizes the speed at which a particle sediments during centrifugation.

<b>Locations, Products, and Sensitivities of the Three Eukaryotic RNA Polymerases</b>			
<b>RNA Polymerase</b>	<b>Cellular Compartment</b>	<b>Product of Transcription</b>	<b><math>\alpha</math>-Amanitin Sensitivity</b>
I	Nucleolus	All rRNAs except 5S rRNA	Insensitive
II	Nucleus	All protein-coding nuclear pre-mRNAs	Extremely sensitive
III	Nucleus	5S rRNA, tRNAs, and small nuclear RNAs	Moderately sensitive

RNA polymerase II is located in the nucleus and synthesizes all protein-coding nuclear pre-mRNAs. Eukaryotic pre-mRNAs undergo extensive processing after transcription but before translation. For clarity, this module’s discussion of transcription and translation in eukaryotes will use the term “mRNAs” to describe only the mature, processed molecules that

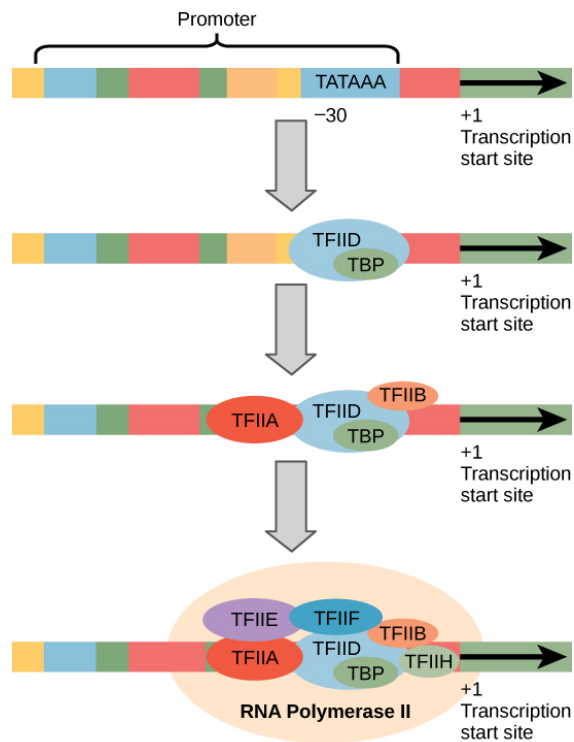
are ready to be translated. RNA polymerase II is responsible for transcribing the overwhelming majority of eukaryotic genes.

RNA polymerase III is also located in the nucleus. This polymerase transcribes a variety of structural RNAs that includes the 5S pre-rRNA, transfer pre-RNAs (pre-tRNAs), and **small nuclear pre-RNAs**. The tRNAs have a critical role in translation; they serve as the adaptor molecules between the mRNA template and the growing polypeptide chain. Small nuclear RNAs have a variety of functions, including “splicing” pre-mRNAs and regulating transcription factors.

A scientist characterizing a new gene can determine which polymerase transcribes it by testing whether the gene is expressed in the presence of a particular mushroom poison,  $\alpha$ -amanitin ([\[link\]](#)). Interestingly,  $\alpha$ -amanitin produced by *Amanita phalloides*, the Death Cap mushroom, affects the three polymerases very differently. RNA polymerase I is completely insensitive to  $\alpha$ -amanitin, meaning that the polymerase can transcribe DNA in vitro in the presence of this poison. In contrast, RNA polymerase II is extremely sensitive to  $\alpha$ -amanitin, and RNA polymerase III is moderately sensitive. Knowing the transcribing polymerase can clue a researcher into the general function of the gene being studied. Because RNA polymerase II transcribes the vast majority of genes, we will focus on this polymerase in our subsequent discussions about eukaryotic transcription factors and promoters.

## Structure of an RNA Polymerase II Promoter

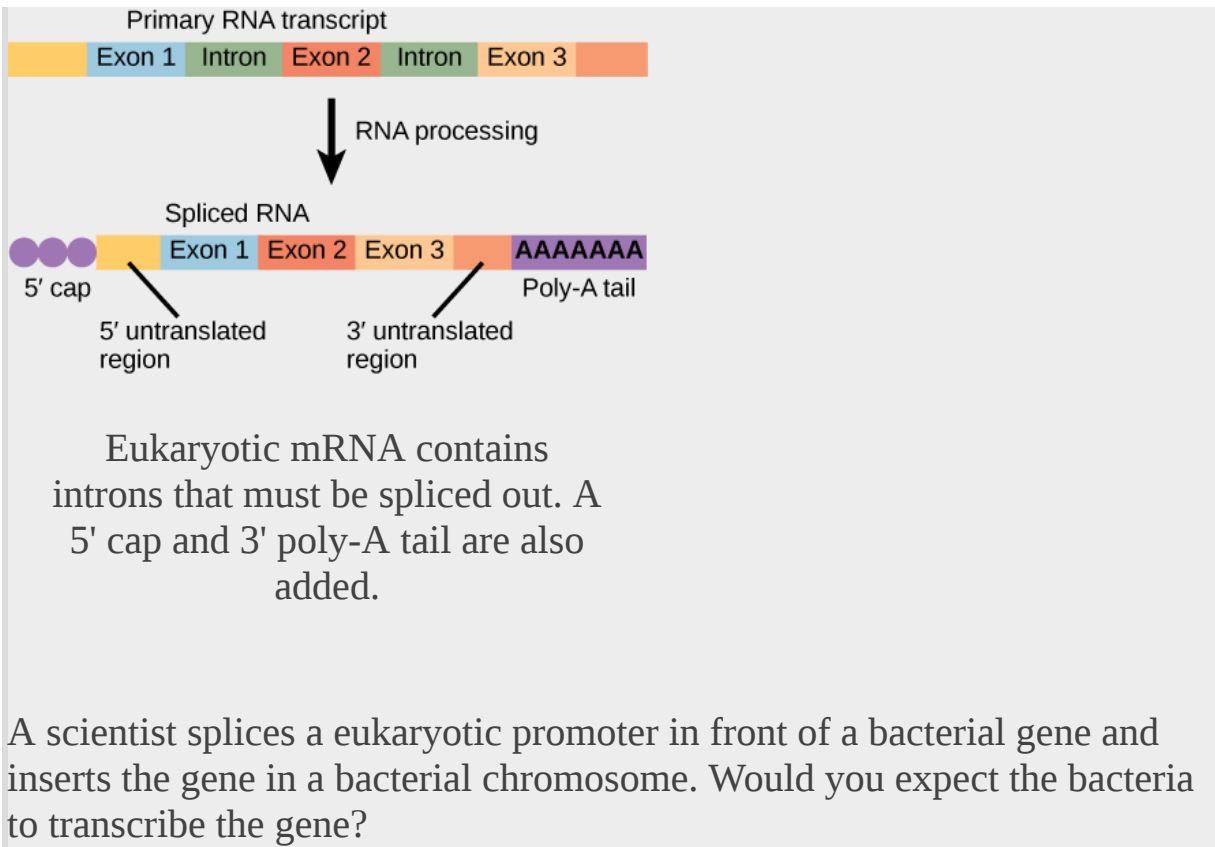
Eukaryotic promoters are much larger and more complex than prokaryotic promoters, but both have a TATA box. For example, in the mouse thymidine kinase gene, the TATA box is located at approximately -30 relative to the initiation (+1) site ([\[link\]](#)). For this gene, the exact TATA box sequence is TATAAAA, as read in the 5' to 3' direction on the nontemplate strand. This sequence is not identical to the *E. coli* TATA box, but it conserves the A–T rich element. The thermostability of A–T bonds is low and this helps the DNA template to locally unwind in preparation for transcription.



A generalized promoter of a gene transcribed by RNA polymerase II is shown.

Transcription factors recognize the promoter. RNA polymerase II then binds and forms the transcription initiation complex.

**Note:**  
Art Connection



The mouse genome includes one gene and two pseudogenes for cytoplasmic thymidine kinase. Pseudogenes are genes that have lost their protein-coding ability or are no longer expressed by the cell. These pseudogenes are copied from mRNA and incorporated into the chromosome. For example, the mouse thymidine kinase promoter also has a conserved **CAAT box** (GGCCAATCT) at approximately -80. This sequence is essential and is involved in binding transcription factors. Further upstream of the TATA box, eukaryotic promoters may also contain one or more **GC-rich boxes** (GGCG) or **octamer boxes** (ATTTGCAT). These elements bind cellular factors that increase the efficiency of transcription initiation and are often identified in more “active” genes that are constantly being expressed by the cell.

## Transcription Factors for RNA Polymerase II



The complexity of eukaryotic transcription does not end with the polymerases and promoters. An army of basal transcription factors, enhancers, and silencers also help to regulate the frequency with which pre-mRNA is synthesized from a gene. Enhancers and silencers affect the efficiency of transcription but are not necessary for transcription to proceed. Basal transcription factors are crucial in the formation of a **preinitiation complex** on the DNA template that subsequently recruits RNA polymerase II for transcription initiation.

The names of the basal transcription factors begin with “TFII” (this is the transcription factor for RNA polymerase II) and are specified with the letters A–J. The transcription factors systematically fall into place on the DNA template, with each one further stabilizing the preinitiation complex and contributing to the recruitment of RNA polymerase II.

The processes of bringing RNA polymerases I and III to the DNA template involve slightly less complex collections of transcription factors, but the general theme is the same. Eukaryotic transcription is a tightly regulated process that requires a variety of proteins to interact with each other and with the DNA strand. Although the process of transcription in eukaryotes involves a greater metabolic investment than in prokaryotes, it ensures that the cell transcribes precisely the pre-mRNAs that it needs for protein synthesis.

### **Note:**

#### **Evolution Connection**

##### **The Evolution of Promoters**

The evolution of genes may be a familiar concept. Mutations can occur in genes during DNA replication, and the result may or may not be beneficial to the cell. By altering an enzyme, structural protein, or some other factor, the process of mutation can transform functions or physical features.

However, eukaryotic promoters and other gene regulatory sequences may evolve as well. For instance, consider a gene that, over many generations, becomes more valuable to the cell. Maybe the gene encodes a structural protein that the cell needs to synthesize in abundance for a certain function. If this is the case, it would be beneficial to the cell for that gene’s promoter

to recruit transcription factors more efficiently and increase gene expression.

Scientists examining the evolution of promoter sequences have reported varying results. In part, this is because it is difficult to infer exactly where a eukaryotic promoter begins and ends. Some promoters occur within genes; others are located very far upstream, or even downstream, of the genes they are regulating. However, when researchers limited their examination to human core promoter sequences that were defined experimentally as sequences that bind the preinitiation complex, they found that promoters evolve even faster than protein-coding genes.

It is still unclear how promoter evolution might correspond to the evolution of humans or other higher organisms. However, the evolution of a promoter to effectively make more or less of a given gene product is an intriguing alternative to the evolution of the genes themselves. [\[footnote\]](#)

H Liang et al., “Fast evolution of core promoters in primate genomes,” *Molecular Biology and Evolution* 25 (2008): 1239–44.

## Promoter Structures for RNA Polymerases I and III

In eukaryotes, the conserved promoter elements differ for genes transcribed by RNA polymerases I, II, and III. RNA polymerase I transcribes genes that have two GC-rich promoter sequences in the -45 to +20 region. These sequences alone are sufficient for transcription initiation to occur, but promoters with additional sequences in the region from -180 to -105 upstream of the initiation site will further enhance initiation. Genes that are transcribed by RNA polymerase III have upstream promoters or promoters that occur within the genes themselves.

## Eukaryotic Elongation and Termination

Following the formation of the preinitiation complex, the polymerase is released from the other transcription factors, and elongation is allowed to proceed as it does in prokaryotes with the polymerase synthesizing pre-mRNA in the 5' to 3' direction. As discussed previously, RNA polymerase

II transcribes the major share of eukaryotic genes, so this section will focus on how this polymerase accomplishes elongation and termination.

Although the enzymatic process of elongation is essentially the same in eukaryotes and prokaryotes, the DNA template is more complex. When eukaryotic cells are not dividing, their genes exist as a diffuse mass of DNA and proteins called chromatin. The DNA is tightly packaged around charged histone proteins at repeated intervals. These DNA–histone complexes, collectively called nucleosomes, are regularly spaced and include 146 nucleotides of DNA wound around eight histones like thread around a spool.

For polynucleotide synthesis to occur, the transcription machinery needs to move histones out of the way every time it encounters a nucleosome. This is accomplished by a special protein complex called **FACT**, which stands for “facilitates chromatin transcription.” This complex pulls histones away from the DNA template as the polymerase moves along it. Once the pre-mRNA is synthesized, the FACT complex replaces the histones to recreate the nucleosomes.

The termination of transcription is different for the different polymerases. Unlike in prokaryotes, elongation by RNA polymerase II in eukaryotes takes place 1,000–2,000 nucleotides beyond the end of the gene being transcribed. This pre-mRNA tail is subsequently removed by cleavage during mRNA processing. On the other hand, RNA polymerases I and III require termination signals. Genes transcribed by RNA polymerase I contain a specific 18-nucleotide sequence that is recognized by a termination protein. The process of termination in RNA polymerase III involves an mRNA hairpin similar to rho-independent termination of transcription in prokaryotes.

## **Section Summary**

Transcription in eukaryotes involves one of three types of polymerases, depending on the gene being transcribed. RNA polymerase II transcribes all of the protein-coding genes, whereas RNA polymerase I transcribes rRNA genes, and RNA polymerase III transcribes rRNA, tRNA, and small nuclear

RNA genes. The initiation of transcription in eukaryotes involves the binding of several transcription factors to complex promoter sequences that are usually located upstream of the gene being copied. The mRNA is synthesized in the 5' to 3' direction, and the FACT complex moves and reassembles nucleosomes as the polymerase passes by. Whereas RNA polymerases I and III terminate transcription by protein- or RNA hairpin-dependent methods, RNA polymerase II transcribes for 1,000 or more nucleotides beyond the gene template and cleaves the excess during pre-mRNA processing.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) A scientist splices a eukaryotic promoter in front of a bacterial gene and inserts the gene in a bacterial chromosome. Would you expect the bacteria to transcribe the gene?

---

#### Solution:

[\[link\]](#) No. Prokaryotes use different promoters than eukaryotes.

## Review Questions

### Exercise:

#### Problem:

Which feature of promoters can be found in both prokaryotes and eukaryotes?

- a. GC box
- b. TATA box
- c. octamer box
- d. -10 and -35 sequences

---

**Solution:**

B

**Exercise:****Problem:**

What transcripts will be most affected by low levels of  $\alpha$ -amanitin?

- a. 18S and 28S rRNAs
- b. pre-mRNAs
- c. 5S rRNAs and tRNAs
- d. other small nuclear RNAs

---

**Solution:**

B

**Glossary****CAAT box**

(GGCCAATCT) essential eukaryotic promoter sequence involved in binding transcription factors

**FACT**

complex that “facilitates chromatin transcription” by disassembling nucleosomes ahead of a transcribing RNA polymerase II and reassembling them after the polymerase passes by

**GC-rich box**

(GGCG) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

**Octamer box**

(ATTTGCAT) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

preinitiation complex

cluster of transcription factors and other proteins that recruit RNA polymerase II for transcription of a DNA template

small nuclear RNA

molecules synthesized by RNA polymerase III that have a variety of functions, including splicing pre-mRNAs and regulating transcription factors

## RNA Processing in Eukaryotes

By the end of this section, you will be able to:

- Describe the different steps in RNA processing
- Understand the significance of exons, introns, and splicing
- Explain how tRNAs and rRNAs are processed

After transcription, eukaryotic pre-mRNAs must undergo several processing steps before they can be translated. Eukaryotic (and prokaryotic) tRNAs and rRNAs also undergo processing before they can function as components in the protein synthesis machinery.

### mRNA Processing

The eukaryotic pre-mRNA undergoes extensive processing before it is ready to be translated. The additional steps involved in eukaryotic mRNA maturation create a molecule with a much longer half-life than a prokaryotic mRNA. Eukaryotic mRNAs last for several hours, whereas the typical *E. coli* mRNA lasts no more than five seconds.

Pre-mRNAs are first coated in RNA-stabilizing proteins; these protect the pre-mRNA from degradation while it is processed and exported out of the nucleus. The three most important steps of pre-mRNA processing are the addition of stabilizing and signaling factors at the 5' and 3' ends of the molecule, and the removal of intervening sequences that do not specify the appropriate amino acids. In rare cases, the mRNA transcript can be “edited” after it is transcribed.

#### **Note:**

##### Evolution Connection

##### **RNA Editing in Trypanosomes**

The trypanosomes are a group of protozoa that include the pathogen *Trypanosoma brucei*, which causes sleeping sickness in humans ([link](#)). Trypanosomes, and virtually all other eukaryotes, have organelles called mitochondria that supply the cell with chemical energy. Mitochondria are

organelles that express their own DNA and are believed to be the remnants of a symbiotic relationship between a eukaryote and an engulfed prokaryote. The mitochondrial DNA of trypanosomes exhibit an interesting exception to The Central Dogma: their pre-mRNAs do not have the correct information to specify a functional protein. Usually, this is because the mRNA is missing several U nucleotides. The cell performs an additional RNA processing step called **RNA editing** to remedy this.



*Trypanosoma brucei* is the causative agent of sleeping sickness in humans. The mRNAs of this pathogen must be modified by the addition of nucleotides before protein synthesis can occur. (credit: modification of work by Torsten Ochsenreiter)

Other genes in the mitochondrial genome encode 40- to 80-nucleotide guide RNAs. One or more of these molecules interacts by complementary base pairing with some of the nucleotides in the pre-mRNA transcript. However, the guide RNA has more A nucleotides than the pre-mRNA has U nucleotides to bind with. In these regions, the guide RNA loops out. The



3' ends of guide RNAs have a long poly-U tail, and these U bases are inserted in regions of the pre-mRNA transcript at which the guide RNAs are looped. This process is entirely mediated by RNA molecules. That is, guide RNAs—rather than proteins—serve as the catalysts in RNA editing. RNA editing is not just a phenomenon of trypanosomes. In the mitochondria of some plants, almost all pre-mRNAs are edited. RNA editing has also been identified in mammals such as rats, rabbits, and even humans. What could be the evolutionary reason for this additional step in pre-mRNA processing? One possibility is that the mitochondria, being remnants of ancient prokaryotes, have an equally ancient RNA-based method for regulating gene expression. In support of this hypothesis, edits made to pre-mRNAs differ depending on cellular conditions. Although speculative, the process of RNA editing may be a holdover from a primordial time when RNA molecules, instead of proteins, were responsible for catalyzing reactions.

## 5' Capping

While the pre-mRNA is still being synthesized, a **7-methylguanosine cap** is added to the 5' end of the growing transcript by a phosphate linkage. This moiety (functional group) protects the nascent mRNA from degradation. In addition, factors involved in protein synthesis recognize the cap to help initiate translation by ribosomes.

## 3' Poly-A Tail

Once elongation is complete, the pre-mRNA is cleaved by an endonuclease between an AAUAAA consensus sequence and a GU-rich sequence, leaving the AAUAAA sequence on the pre-mRNA. An enzyme called poly-A polymerase then adds a string of approximately 200 A residues, called the **poly-A tail**. This modification further protects the pre-mRNA from degradation and signals the export of the cellular factors that the transcript needs to the cytoplasm.

## Pre-mRNA Splicing

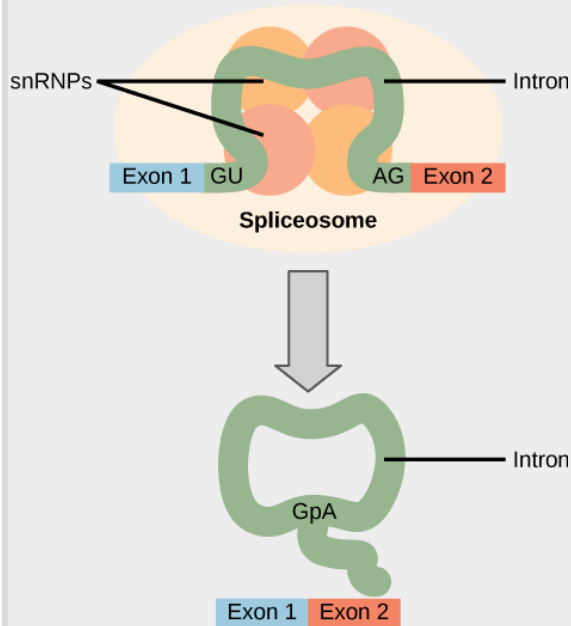
Eukaryotic genes are composed of **exons**, which correspond to protein-coding sequences (*ex-on* signifies that they are *expressed*), and *intervening* sequences called **introns** (*int-ron* denotes their *intervening* role), which may be involved in gene regulation but are removed from the pre-mRNA during processing. Intron sequences in mRNA do not encode functional proteins.

The discovery of introns came as a surprise to researchers in the 1970s who expected that pre-mRNAs would specify protein sequences without further processing, as they had observed in prokaryotes. The genes of higher eukaryotes very often contain one or more introns. These regions may correspond to regulatory sequences; however, the biological significance of having many introns or having very long introns in a gene is unclear. It is possible that introns slow down gene expression because it takes longer to transcribe pre-mRNAs with lots of introns. Alternatively, introns may be nonfunctional sequence remnants left over from the fusion of ancient genes throughout evolution. This is supported by the fact that separate exons often encode separate protein subunits or domains. For the most part, the sequences of introns can be mutated without ultimately affecting the protein product.

All of a pre-mRNA's introns must be completely and precisely removed before protein synthesis. If the process errs by even a single nucleotide, the reading frame of the rejoined exons would shift, and the resulting protein would be dysfunctional. The process of removing introns and reconnecting exons is called **splicing** ([\[link\]](#)). Introns are removed and degraded while the pre-mRNA is still in the nucleus. Splicing occurs by a sequence-specific mechanism that ensures introns will be removed and exons rejoined with the accuracy and precision of a single nucleotide. The splicing of pre-mRNAs is conducted by complexes of proteins and RNA molecules called spliceosomes.

### Note:

Art Connection



Pre-mRNA splicing involves the precise removal of introns from the primary RNA transcript. The splicing process is catalyzed by protein complexes called spliceosomes that are composed of proteins and RNA molecules called snRNAs. Spliceosomes recognize sequences at the 5' and 3' end of the intron.

Errors in splicing are implicated in cancers and other human diseases. What kinds of mutations might lead to splicing errors? Think of different possible outcomes if splicing errors occur.

Note that more than 70 individual introns can be present, and each has to undergo the process of splicing—in addition to 5' capping and the addition

of a poly-A tail—just to generate a single, translatable mRNA molecule.

**Note:**

Link to Learning



See how introns are removed during RNA splicing [at this website](#).

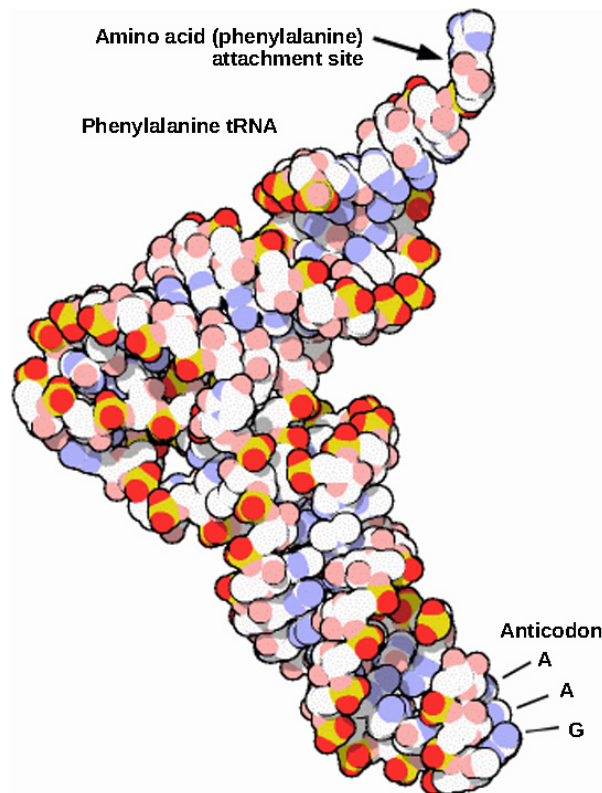
## Processing of tRNAs and rRNAs

The tRNAs and rRNAs are structural molecules that have roles in protein synthesis; however, these RNAs are not themselves translated. Pre-rRNAs are transcribed, processed, and assembled into ribosomes in the nucleolus. Pre-tRNAs are transcribed and processed in the nucleus and then released into the cytoplasm where they are linked to free amino acids for protein synthesis.

Most of the tRNAs and rRNAs in eukaryotes and prokaryotes are first transcribed as a long precursor molecule that spans multiple rRNAs or tRNAs. Enzymes then cleave the precursors into subunits corresponding to each structural RNA. Some of the bases of pre-rRNAs are methylated; that is, a  $\text{-CH}_3$  moiety (methyl functional group) is added for stability. Pre-tRNA molecules also undergo methylation. As with pre-mRNAs, subunit excision occurs in eukaryotic pre-RNAs destined to become tRNAs or rRNAs.

Mature rRNAs make up approximately 50 percent of each ribosome. Some of a ribosome's RNA molecules are purely structural, whereas others have catalytic or binding activities. Mature tRNAs take on a three-dimensional

structure through intramolecular hydrogen bonding to position the amino acid binding site at one end and the **anticodon** at the other end ([\[link\]](#)). The anticodon is a three-nucleotide sequence in a tRNA that interacts with an mRNA codon through complementary base pairing.



This is a space-filling model of a tRNA molecule that adds the amino acid phenylalanine to a growing polypeptide chain. The anticodon AAG binds the Codon UUC on the mRNA. The amino acid phenylalanine is attached to the other end of the tRNA.

## Section Summary

Eukaryotic pre-mRNAs are modified with a 5' methylguanosine cap and a poly-A tail. These structures protect the mature mRNA from degradation and help export it from the nucleus. Pre-mRNAs also undergo splicing, in which introns are removed and exons are reconnected with single-nucleotide accuracy. Only finished mRNAs that have undergone 5' capping, 3' polyadenylation, and intron splicing are exported from the nucleus to the cytoplasm. Pre-rRNAs and pre-tRNAs may be processed by intramolecular cleavage, splicing, methylation, and chemical conversion of nucleotides. Rarely, RNA editing is also performed to insert missing bases after an mRNA has been synthesized.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Errors in splicing are implicated in cancers and other human diseases. What kinds of mutations might lead to splicing errors? Think of different possible outcomes if splicing errors occur.

---

#### Solution:

[\[link\]](#) Mutations in the spliceosome recognition sequence at each end of the intron, or in the proteins and RNAs that make up the spliceosome, may impair splicing. Mutations may also add new spliceosome recognition sites. Splicing errors could lead to introns being retained in spliced RNA, exons being excised, or changes in the location of the splice site.

## Review Questions

### Exercise:

**Problem:**

Which pre-mRNA processing step is important for initiating translation?

- a. poly-A tail
- b. RNA editing
- c. splicing
- d. 7-methylguanosine cap

---

**Solution:**

D

**Exercise:****Problem:**

What processing step enhances the stability of pre-tRNAs and pre-rRNAs?

- a. methylation
- b. nucleotide modification
- c. cleavage
- d. splicing

---

**Solution:**

A

**Glossary**

7-methylguanosine cap

modification added to the 5' end of pre-mRNAs to protect mRNA from degradation and assist translation

anticodon

three-nucleotide sequence in a tRNA molecule that corresponds to an mRNA codon

exon

sequence present in protein-coding mRNA after completion of pre-mRNA splicing

intron

non-protein-coding intervening sequences that are spliced from mRNA during processing

poly-A tail

modification added to the 3' end of pre-mRNAs to protect mRNA from degradation and assist mRNA export from the nucleus

RNA editing

direct alteration of one or more nucleotides in an mRNA that has already been synthesized

splicing

process of removing introns and reconnecting exons in a pre-mRNA

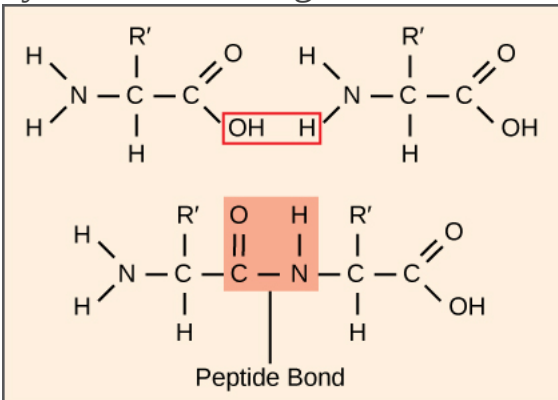


## Ribosomes and Protein Synthesis

By the end of this section, you will be able to:

- Describe the different steps in protein synthesis
- Discuss the role of ribosomes in protein synthesis

The synthesis of proteins consumes more of a cell's energy than any other metabolic process. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform virtually every function of a cell. The process of translation, or protein synthesis, involves the decoding of an mRNA message into a polypeptide product. Amino acids are covalently strung together by interlinking peptide bonds in lengths ranging from approximately 50 amino acid residues to more than 1,000. Each individual amino acid has an amino group ( $\text{NH}_2$ ) and a carboxyl ( $\text{COOH}$ ) group. Polypeptides are formed when the amino group of one amino acid forms an amide (i.e., peptide) bond with the carboxyl group of another amino acid ([\[link\]](#)). This reaction is catalyzed by ribosomes and generates one water molecule.



A peptide bond links the carboxyl end of one amino acid with the amino end of another, expelling one water molecule. For simplicity in this image, only the functional groups involved in the peptide bond are shown. The R and R' designations

refer to the rest of each amino acid structure.

## The Protein Synthesis Machinery

In addition to the mRNA template, many molecules and macromolecules contribute to the process of translation. The composition of each component may vary across species; for instance, ribosomes may consist of different numbers of rRNAs and polypeptides depending on the organism. However, the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, ribosomes, tRNAs, and various enzymatic factors.

### Note:

#### Link to Learning



Click through the steps of this [PBS interactive](#) to see protein synthesis in action.

## Ribosomes

Even before an mRNA is translated, a cell must invest energy to build each of its ribosomes. In *E. coli*, there are between 10,000 and 70,000 ribosomes present in each cell at any given time. A ribosome is a complex macromolecule composed of structural and catalytic rRNAs, and many distinct polypeptides. In eukaryotes, the nucleolus is completely specialized for the synthesis and assembly of rRNAs.

Ribosomes exist in the cytoplasm in prokaryotes and in the cytoplasm and rough endoplasmic reticulum in eukaryotes. Mitochondria and chloroplasts also have their own ribosomes in the matrix and stroma, which look more similar to prokaryotic ribosomes (and have similar drug sensitivities) than the ribosomes just outside their outer membranes in the cytoplasm.

Ribosomes dissociate into large and small subunits when they are not synthesizing proteins and reassociate during the initiation of translation. In *E. coli*, the small subunit is described as 30S, and the large subunit is 50S, for a total of 70S (recall that Svedberg units are not additive). Mammalian ribosomes have a small 40S subunit and a large 60S subunit, for a total of 80S. The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds tRNAs. Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction: reading the mRNA from 5' to 3' and synthesizing the polypeptide from the N terminus to the C terminus. The complete mRNA/poly-ribosome structure is called a **polysome**.

## **tRNAs**

The tRNAs are structural RNA molecules that were transcribed from genes by RNA polymerase III. Depending on the species, 40 to 60 types of tRNAs exist in the cytoplasm. Serving as adaptors, specific tRNAs bind to sequences on the mRNA template and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually “translate” the language of RNA into the language of proteins.

Of the 64 possible mRNA codons—or triplet combinations of A, U, G, and C—three specify the termination of protein synthesis and 61 specify the addition of amino acids to the polypeptide chain. Of these 61, one codon (AUG) also encodes the initiation of translation. Each tRNA anticodon can base pair with one of the mRNA codons and add an amino acid or terminate translation, according to the genetic code. For instance, if the sequence CUA occurred on an mRNA template in the proper reading frame, it would bind a tRNA expressing the complementary sequence, GAU, which would be linked to the amino acid leucine.

As the adaptor molecules of translation, it is surprising that tRNAs can fit so much specificity into such a small package. Consider that tRNAs need to interact with three factors: 1) they must be recognized by the correct aminoacyl synthetase (see below); 2) they must be recognized by ribosomes; and 3) they must bind to the correct sequence in mRNA.

## Aminoacyl tRNA Synthetases

The process of pre-tRNA synthesis by RNA polymerase III only creates the RNA portion of the adaptor molecule. The corresponding amino acid must be added later, once the tRNA is processed and exported to the cytoplasm. Through the process of tRNA “charging,” each tRNA molecule is linked to its correct amino acid by a group of enzymes called **aminoacyl tRNA synthetases**. At least one type of aminoacyl tRNA synthetase exists for each of the 20 amino acids; the exact number of aminoacyl tRNA synthetases varies by species. These enzymes first bind and hydrolyze ATP to catalyze a high-energy bond between an amino acid and adenosine monophosphate (AMP); a pyrophosphate molecule is expelled in this reaction. The activated amino acid is then transferred to the tRNA, and AMP is released.

## The Mechanism of Protein Synthesis

As with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in prokaryotes and eukaryotes. Here we’ll explore how translation occurs in *E. coli*, a representative prokaryote, and specify any differences between prokaryotic and eukaryotic translation.

### Initiation of Translation

Protein synthesis begins with the formation of an initiation complex. In *E. coli*, this complex involves the small 30S ribosome, the mRNA template, three initiation factors (IFs; IF-1, IF-2, and IF-3), and a special **initiator**

**tRNA**, called  $\text{tRNA}_f^{\text{Met}}$ . The initiator tRNA interacts with the **start codon** AUG (or rarely, GUG), links to a formylated methionine called fMet, and can also bind IF-2. Formylated methionine is inserted by  $\text{fMet} - \text{tRNA}_f^{\text{Met}}$  at the beginning of every polypeptide chain synthesized by *E. coli*, but it is usually clipped off after translation is complete. When an in-frame AUG is encountered during translation elongation, a non-formylated methionine is inserted by a regular  $\text{Met-tRNA}^{\text{Met}}$ .

In *E. coli* mRNA, a sequence upstream of the first AUG codon, called the **Shine-Dalgarno sequence** (AGGAGG), interacts with the rRNA molecules that compose the ribosome. This interaction anchors the 30S ribosomal subunit at the correct location on the mRNA template. Guanosine triphosphate (GTP), which is a purine nucleotide triphosphate, acts as an energy source during translation—both at the start of elongation and during the ribosome's translocation.

In eukaryotes, a similar initiation complex forms, comprising mRNA, the 40S small ribosomal subunit, IFs, and nucleoside triphosphates (GTP and ATP). The charged initiator tRNA, called  $\text{Met-tRNA}_i$ , does not bind fMet in eukaryotes, but is distinct from other Met-tRNAs in that it can bind IFs.

Instead of depositing at the Shine-Dalgarno sequence, the eukaryotic initiation complex recognizes the 7-methylguanosine cap at the 5' end of the mRNA. A cap-binding protein (CBP) and several other IFs assist the movement of the ribosome to the 5' cap. Once at the cap, the initiation complex tracks along the mRNA in the 5' to 3' direction, searching for the AUG start codon. Many eukaryotic mRNAs are translated from the first AUG, but this is not always the case. According to **Kozak's rules**, the nucleotides around the AUG indicate whether it is the correct start codon. Kozak's rules state that the following consensus sequence must appear around the AUG of vertebrate genes: 5'-gccRccAUGG-3'. The R (for purine) indicates a site that can be either A or G, but cannot be C or U. Essentially, the closer the sequence is to this consensus, the higher the efficiency of translation.

Once the appropriate AUG is identified, the other proteins and CBP dissociate, and the 60S subunit binds to the complex of  $\text{Met-tRNA}_i$ , mRNA,

and the 40S subunit. This step completes the initiation of translation in eukaryotes.

## Translation, Elongation, and Termination

In prokaryotes and eukaryotes, the basics of elongation are the same, so we will review elongation from the perspective of *E. coli*. The 50S ribosomal subunit of *E. coli* consists of three compartments: the A (aminoacyl) site binds incoming charged aminoacyl tRNAs. The P (peptidyl) site binds charged tRNAs carrying amino acids that have formed peptide bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA. The E (exit) site releases dissociated tRNAs so that they can be recharged with free amino acids. There is one exception to this assembly line of tRNAs: in *E. coli*, fMet – tRNA<sub>f</sub><sup>Met</sup> is capable of entering the P site directly without first entering the A site. Similarly, the eukaryotic Met-tRNA<sub>i</sub>, with help from other proteins of the initiation complex, binds directly to the P site. In both cases, this creates an initiation complex with a free A site ready to accept the tRNA corresponding to the first codon after the AUG.

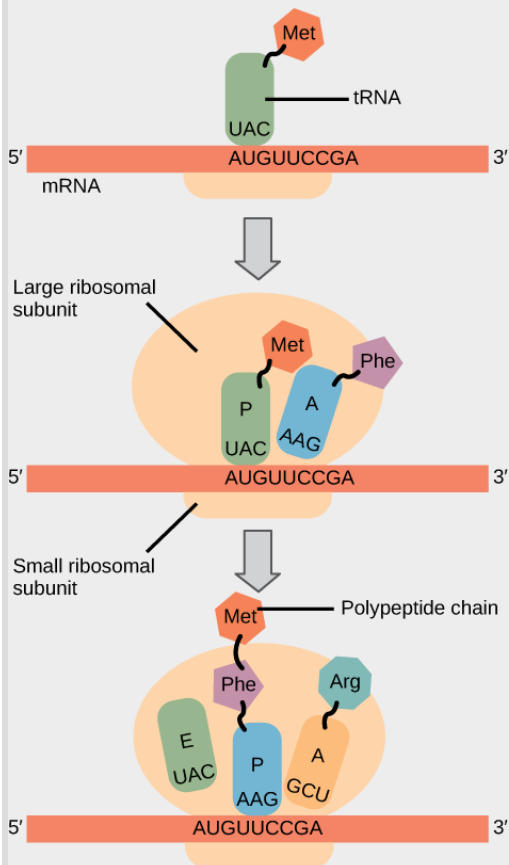
During translation elongation, the mRNA template provides specificity. As the ribosome moves along the mRNA, each mRNA codon comes into register, and specific binding with the corresponding charged tRNA anticodon is ensured. If mRNA were not present in the elongation complex, the ribosome would bind tRNAs nonspecifically.

Elongation proceeds with charged tRNAs entering the A site and then shifting to the P site followed by the E site with each single-codon “step” of the ribosome. Ribosomal steps are induced by conformational changes that advance the ribosome by three bases in the 3' direction. The energy for each step of the ribosome is donated by an elongation factor that hydrolyzes GTP. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by **peptidyl transferase**, an RNA-based enzyme that is integrated into the 50S ribosomal subunit. The energy for each peptide bond formation

is derived from GTP hydrolysis, which is catalyzed by a separate elongation factor. The amino acid bound to the P-site tRNA is also linked to the growing polypeptide chain. As the ribosome steps across the mRNA, the former P-site tRNA enters the E site, detaches from the amino acid, and is expelled ([link](#)). Amazingly, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200-amino acid protein can be translated in just 10 seconds.

### Note:

#### Art Connection



Translation begins when an initiator tRNA anticodon recognizes a codon on mRNA. The large ribosomal subunit

joins the small subunit,  
and a second tRNA is  
recruited. As the mRNA  
moves relative to the  
ribosome, the polypeptide  
chain is formed. Entry of  
a release factor into the A  
site terminates translation  
and the components  
dissociate.

Many antibiotics inhibit bacterial protein synthesis. For example, tetracycline blocks the A site on the bacterial ribosome, and chloramphenicol blocks peptidyl transfer. What specific effect would you expect each of these antibiotics to have on protein synthesis?

Tetracycline would directly affect:

- a. tRNA binding to the ribosome
- b. ribosome assembly
- c. growth of the protein chain

Chloramphenicol would directly affect

- a. tRNA binding to the ribosome
- b. ribosome assembly
- c. growth of the protein chain

Termination of translation occurs when a nonsense codon (UAA, UAG, or UGA) is encountered. Upon aligning with the A site, these nonsense codons are recognized by release factors in prokaryotes and eukaryotes that instruct peptidyl transferase to add a water molecule to the carboxyl end of the P-site amino acid. This reaction forces the P-site amino acid to detach from its tRNA, and the newly made protein is released. The small and large ribosomal subunits dissociate from the mRNA and from each other; they



are recruited almost immediately into another translation initiation complex. After many ribosomes have completed translation, the mRNA is degraded so the nucleotides can be reused in another transcription reaction.

## **Protein Folding, Modification, and Targeting**

During and after translation, individual amino acids may be chemically modified, signal sequences may be appended, and the new protein “folds” into a distinct three-dimensional structure as a result of intramolecular interactions. A **signal sequence** is a short tail of amino acids that directs a protein to a specific cellular compartment. These sequences at the amino end or the carboxyl end of the protein can be thought of as the protein’s “train ticket” to its ultimate destination. Other cellular factors recognize each signal sequence and help transport the protein from the cytoplasm to its correct compartment. For instance, a specific sequence at the amino terminus will direct a protein to the mitochondria or chloroplasts (in plants). Once the protein reaches its cellular destination, the signal sequence is usually clipped off.

Many proteins fold spontaneously, but some proteins require helper molecules, called chaperones, to prevent them from aggregating during the complicated process of folding. Even if a protein is properly specified by its corresponding mRNA, it could take on a completely dysfunctional shape if abnormal temperature or pH conditions prevent it from folding correctly.

## **Section Summary**

The players in translation include the mRNA template, ribosomes, tRNAs, and various enzymatic factors. The small ribosomal subunit forms on the mRNA template either at the Shine-Dalgarno sequence (prokaryotes) or the 5' cap (eukaryotes). Translation begins at the initiating AUG on the mRNA, specifying methionine. The formation of peptide bonds occurs between sequential amino acids specified by the mRNA template according to the genetic code. Charged tRNAs enter the ribosomal A site, and their amino acid bonds with the amino acid at the P site. The entire mRNA is translated in three-nucleotide “steps” of the ribosome. When a nonsense codon is encountered, a release factor binds and dissociates the components and

frees the new protein. Folding of the protein occurs during and after translation.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Many antibiotics inhibit bacterial protein synthesis. For example, tetracycline blocks the A site on the bacterial ribosome, and chloramphenicol blocks peptidyl transfer. What specific effect would you expect each of these antibiotics to have on protein synthesis?

Tetracycline would directly affect:

- a. tRNA binding to the ribosome
- b. ribosome assembly
- c. growth of the protein chain

Chloramphenicol would directly affect

- a. tRNA binding to the ribosome
- b. ribosome assembly
- c. growth of the protein chain

---

#### Solution:

[\[link\]](#) Tetracycline: a; Chloramphenicol: c.

## Review Questions

### Exercise:

#### Problem:

The RNA components of ribosomes are synthesized in the \_\_\_\_\_.

- a. cytoplasm
  - b. nucleus
  - c. nucleolus
  - d. endoplasmic reticulum
- 

**Solution:**

C

**Exercise:**

**Problem:**

In any given species, there are at least how many types of aminoacyl tRNA synthetases?

- a. 20
  - b. 40
  - c. 100
  - d. 200
- 

**Solution:**

A

**Free Response**

**Exercise:**

**Problem:**

Transcribe and translate the following DNA sequence (nontemplate strand): 5'-ATGGCCGGTTATTAAGCA-3'

---

**Solution:**

The mRNA would be: 5'-AUGGCCGGUUAUUAAGCA-3'. The protein would be: MAGY. Even though there are six codons, the fifth codon corresponds to a stop, so the sixth codon would not be translated.

### Exercise:

#### Problem:

Explain how single nucleotide changes can have vastly different effects on protein function.

---

#### Solution:

Nucleotide changes in the third position of codons may not change the amino acid and would have no effect on the protein. Other nucleotide changes that change important amino acids or create or delete start or stop codons would have severe effects on the amino acid sequence of the protein.

## Glossary

aminoacyl tRNA synthetase

enzyme that “charges” tRNA molecules by catalyzing a bond between the tRNA and a corresponding amino acid

initiator tRNA

in prokaryotes, called  $tRNA_f^{Met}$ ; in eukaryotes, called tRNA<sub>i</sub>; a tRNA that interacts with a start codon, binds directly to the ribosome P site, and links to a special methionine to begin a polypeptide chain

Kozak’s rules

determines the correct initiation AUG in a eukaryotic mRNA; the following consensus sequence must appear around the AUG: 5'-GCC(**purine**)CCA**AUG**G-3'; the bolded bases are most important

peptidyl transferase

RNA-based enzyme that is integrated into the 50S ribosomal subunit and catalyzes the formation of peptide bonds

polysome

mRNA molecule simultaneously being translated by many ribosomes all going in the same direction

Shine-Dalgarno sequence

(AGGAGG); initiates prokaryotic translation by interacting with rRNA molecules comprising the 30S ribosome

signal sequence

short tail of amino acids that directs a protein to a specific cellular compartment

start codon

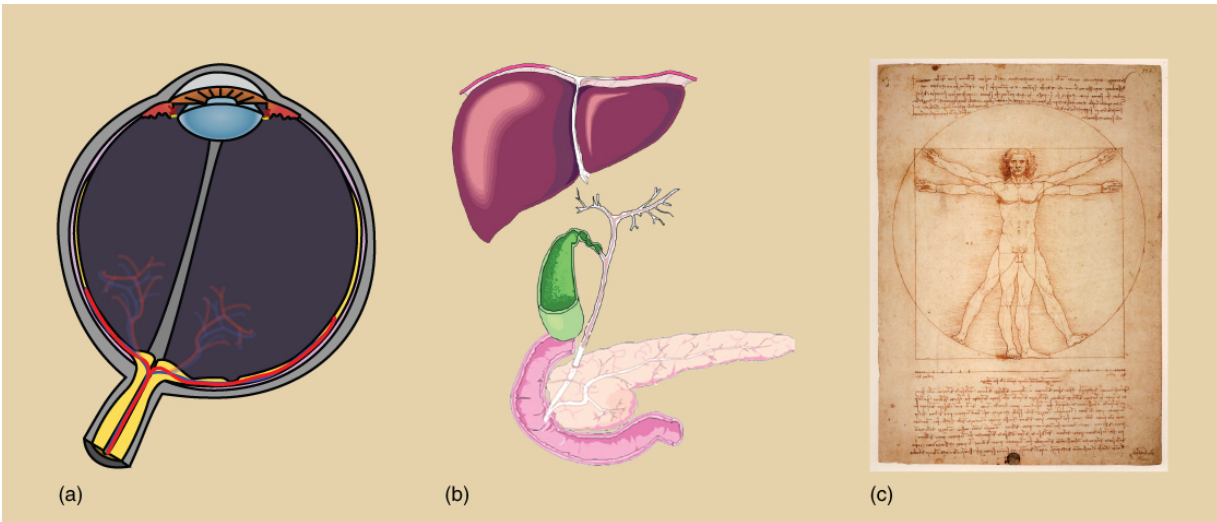
AUG (or rarely, GUG) on an mRNA from which translation begins; always specifies methionine

## Introduction

class="introduction"

The genetic content of each somatic cell in an organism is the same, but not all genes are expressed in every cell. The control of which genes are expressed dictates whether a cell is (a) an eye cell or (b) a liver cell. It is the differential gene expression patterns that arise in different cells that give rise

to (c) a complete organism.



Each somatic cell in the body generally contains the same DNA. A few exceptions include red blood cells, which contain no DNA in their mature state, and some immune system cells that rearrange their DNA while producing antibodies. In general, however, the genes that determine whether you have green eyes, brown hair, and how fast you metabolize food are the same in the cells in your eyes and your liver, even though these organs function quite differently. If each cell has the same DNA, how is it that cells or organs are different? Why do cells in the eye differ so dramatically from cells in the liver?

Whereas each cell shares the same genome and DNA sequence, each cell does not turn on, or express, the same set of genes. Each cell type needs a different set of proteins to perform its function. Therefore, only a small subset of proteins is expressed in a cell. For the proteins to be expressed, the DNA must be transcribed into RNA and the RNA must be translated into protein. In a given cell type, not all genes encoded in the DNA are transcribed into RNA or translated into protein because specific cells in our body have specific functions. Specialized proteins that make up the eye (iris, lens, and cornea) are only expressed in the eye, whereas the specialized proteins in the heart (pacemaker cells, heart muscle, and valves)

are only expressed in the heart. At any given time, only a subset of all of the genes encoded by our DNA are expressed and translated into proteins. The expression of specific genes is a highly regulated process with many levels and stages of control. This complexity ensures the proper expression in the proper cell at the proper time.



## Regulation of Gene Expression

By the end of this section, you will be able to:

- Discuss why every cell does not express all of its genes
- Describe how prokaryotic gene regulation occurs at the transcriptional level
- Discuss how eukaryotic gene regulation occurs at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels

For a cell to function properly, necessary proteins must be synthesized at the proper time. All cells control or regulate the synthesis of proteins from information encoded in their DNA. The process of turning on a gene to produce RNA and protein is called **gene expression**. Whether in a simple unicellular organism or a complex multi-cellular organism, each cell controls when and how its genes are expressed. For this to occur, there must be a mechanism to control when a gene is expressed to make RNA and protein, how much of the protein is made, and when it is time to stop making that protein because it is no longer needed.

The regulation of gene expression conserves energy and space. It would require a significant amount of energy for an organism to express every gene at all times, so it is more energy efficient to turn on the genes only when they are required. In addition, only expressing a subset of genes in each cell saves space because DNA must be unwound from its tightly coiled structure to transcribe and translate the DNA. Cells would have to be enormous if every protein were expressed in every cell all the time.

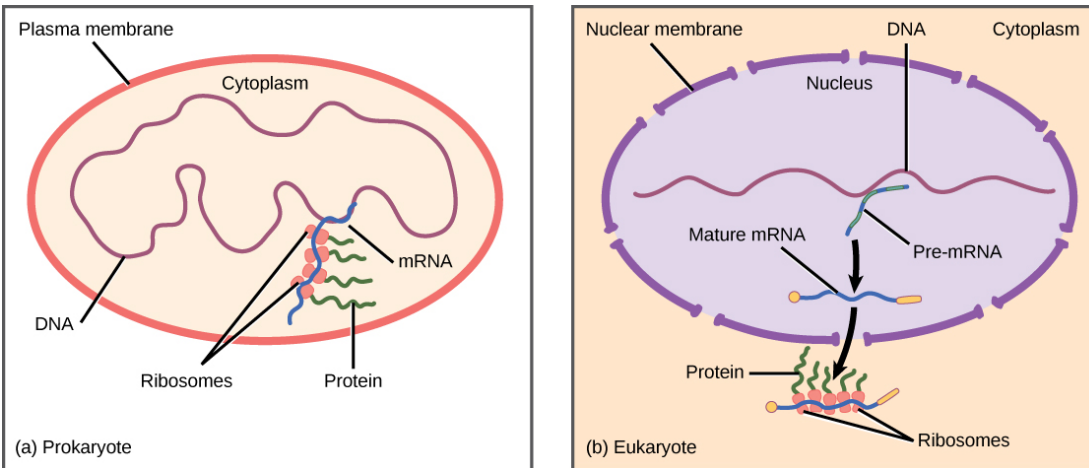
The control of gene expression is extremely complex. Malfunctions in this process are detrimental to the cell and can lead to the development of many diseases, including cancer.

## Prokaryotic versus Eukaryotic Gene Expression

To understand how gene expression is regulated, we must first understand how a gene codes for a functional protein in a cell. The process occurs in both prokaryotic and eukaryotic cells, just in slightly different manners.

Prokaryotic organisms are single-celled organisms that lack a cell nucleus, and their DNA therefore floats freely in the cell cytoplasm. To synthesize a protein, the processes of transcription and translation occur almost simultaneously. When the resulting protein is no longer needed, transcription stops. As a result, the primary method to control what type of protein and how much of each protein is expressed in a prokaryotic cell is the regulation of DNA transcription. All of the subsequent steps occur automatically. When more protein is required, more transcription occurs. Therefore, in prokaryotic cells, the control of gene expression is mostly at the transcriptional level.

Eukaryotic cells, in contrast, have intracellular organelles that add to their complexity. In eukaryotic cells, the DNA is contained inside the cell's nucleus and there it is transcribed into RNA. The newly synthesized RNA is then transported out of the nucleus into the cytoplasm, where ribosomes translate the RNA into protein. The processes of transcription and translation are physically separated by the nuclear membrane; transcription occurs only within the nucleus, and translation occurs only outside the nucleus in the cytoplasm. The regulation of gene expression can occur at all stages of the process ([\[link\]](#)). Regulation may occur when the DNA is uncoiled and loosened from nucleosomes to bind transcription factors (**epigenetic** level), when the RNA is transcribed (transcriptional level), when the RNA is processed and exported to the cytoplasm after it is transcribed (**post-transcriptional** level), when the RNA is translated into protein (translational level), or after the protein has been made (**post-translational** level).



Prokaryotic transcription and translation occur simultaneously in the cytoplasm, and regulation occurs at the transcriptional level. Eukaryotic gene expression is regulated during transcription and RNA processing, which take place in the nucleus, and during protein translation, which takes place in the cytoplasm. Further regulation may occur through post-translational modifications of proteins.

The differences in the regulation of gene expression between prokaryotes and eukaryotes are summarized in [\[link\]](#). The regulation of gene expression is discussed in detail in subsequent modules.

<b>Differences in the Regulation of Gene Expression of Prokaryotic and Eukaryotic Organisms</b>	
<b>Prokaryotic organisms</b>	<b>Eukaryotic organisms</b>
Lack nucleus	Contain nucleus

<b>Differences in the Regulation of Gene Expression of Prokaryotic and Eukaryotic Organisms</b>	
<b>Prokaryotic organisms</b>	<b>Eukaryotic organisms</b>
DNA is found in the cytoplasm	DNA is confined to the nuclear compartment
RNA transcription and protein formation occur almost simultaneously	RNA transcription occurs prior to protein formation, and it takes place in the nucleus. Translation of RNA to protein occurs in the cytoplasm.
Gene expression is regulated primarily at the transcriptional level	Gene expression is regulated at many levels (epigenetic, transcriptional, nuclear shuttling, post-transcriptional, translational, and post-translational)

**Note:**

**Evolution Connection**

**Evolution of Gene Regulation**

Prokaryotic cells can only regulate gene expression by controlling the amount of transcription. As eukaryotic cells evolved, the complexity of the control of gene expression increased. For example, with the evolution of eukaryotic cells came compartmentalization of important cellular components and cellular processes. A nuclear region that contains the DNA was formed. Transcription and translation were physically separated into two different cellular compartments. It therefore became possible to control gene expression by regulating transcription in the nucleus, and also by controlling the RNA levels and protein translation present outside the nucleus.

Some cellular processes arose from the need of the organism to defend itself. Cellular processes such as gene silencing developed to protect the cell from viral or parasitic infections. If the cell could quickly shut off gene expression for a short period of time, it would be able to survive an infection when other organisms could not. Therefore, the organism evolved a new process that helped it survive, and it was able to pass this new development to offspring.

## Section Summary

While all somatic cells within an organism contain the same DNA, not all cells within that organism express the same proteins. Prokaryotic organisms express the entire DNA they encode in every cell, but not necessarily all at the same time. Proteins are expressed only when they are needed.

Eukaryotic organisms express a subset of the DNA that is encoded in any given cell. In each cell type, the type and amount of protein is regulated by controlling gene expression. To express a protein, the DNA is first transcribed into RNA, which is then translated into proteins. In prokaryotic cells, these processes occur almost simultaneously. In eukaryotic cells, transcription occurs in the nucleus and is separate from the translation that occurs in the cytoplasm. Gene expression in prokaryotes is mostly regulated at the transcriptional level (some epigenetic and post-translational regulation is also present), whereas in eukaryotic cells, gene expression is regulated at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels.

## Review Questions

### Exercise:

#### Problem:

Control of gene expression in eukaryotic cells occurs at which level(s)?

- a. only the transcriptional level

- b. epigenetic and transcriptional levels
  - c. epigenetic, transcriptional, and translational levels
  - d. epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels
- 

**Solution:**

D

**Exercise:**

**Problem:** Post-translational control refers to:

- a. regulation of gene expression after transcription
  - b. regulation of gene expression after translation
  - c. control of epigenetic activation
  - d. period between transcription and translation
- 

**Solution:**

B

**Free Response**

**Exercise:**

**Problem:**

Name two differences between prokaryotic and eukaryotic cells and how these differences benefit multicellular organisms.

---

**Solution:**

Eukaryotic cells have a nucleus, whereas prokaryotic cells do not. In eukaryotic cells, DNA is confined within the nuclear region. Because of this, transcription and translation are physically separated. This

creates a more complex mechanism for the control of gene expression that benefits multicellular organisms because it compartmentalizes gene regulation.

Gene expression occurs at many stages in eukaryotic cells, whereas in prokaryotic cells, control of gene expression only occurs at the transcriptional level. This allows for greater control of gene expression in eukaryotes and more complex systems to be developed. Because of this, different cell types can arise in an individual organism.

### **Exercise:**

#### **Problem:**

Describe how controlling gene expression will alter the overall protein levels in the cell.

---

#### **Solution:**

The cell controls which proteins are expressed and to what level each protein is expressed in the cell. Prokaryotic cells alter the transcription rate to turn genes on or off. This method will increase or decrease protein levels in response to what is needed by the cell. Eukaryotic cells change the accessibility (epigenetic), transcription, or translation of a gene. This will alter the amount of RNA and the lifespan of the RNA to alter the amount of protein that exists. Eukaryotic cells also control protein translation to increase or decrease the overall levels. Eukaryotic organisms are much more complex and can manipulate protein levels by changing many stages in the process.

### **Glossary**

epigenetic

heritable changes that do not involve changes in the DNA sequence

gene expression

processes that control the turning on or turning off of a gene

post-transcriptional

control of gene expression after the RNA molecule has been created  
but before it is translated into protein

post-translational

control of gene expression after a protein has been created



## Prokaryotic Gene Regulation

By the end of this section, you will be able to:

- Describe the steps involved in prokaryotic gene regulation
- Explain the roles of activators, inducers, and repressors in gene regulation

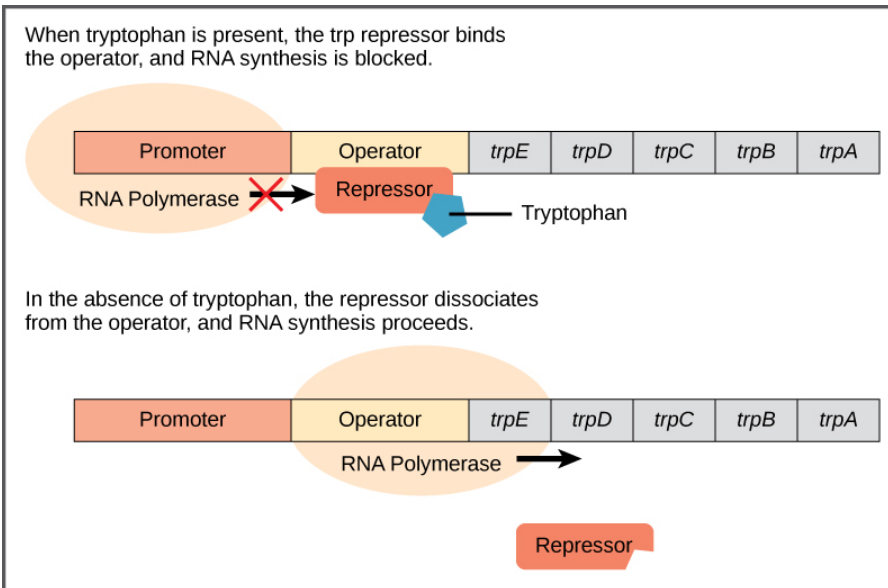
The DNA of prokaryotes is organized into a circular chromosome supercoiled in the nucleoid region of the cell cytoplasm. Proteins that are needed for a specific function, or that are involved in the same biochemical pathway, are encoded together in blocks called **operons**. For example, all of the genes needed to use lactose as an energy source are coded next to each other in the lactose (or *lac*) operon.

In prokaryotic cells, there are three types of regulatory molecules that can affect the expression of operons: repressors, activators, and inducers.

**Repressors** are proteins that suppress transcription of a gene in response to an external stimulus, whereas **activators** are proteins that increase the transcription of a gene in response to an external stimulus. Finally, inducers are small molecules that either activate or repress transcription depending on the needs of the cell and the availability of substrate.

### The *trp* Operon: A Repressor Operon

Bacteria such as *E. coli* need amino acids to survive. **Tryptophan** is one such amino acid that *E. coli* can ingest from the environment. *E. coli* can also synthesize tryptophan using enzymes that are encoded by five genes. These five genes are next to each other in what is called the **tryptophan (*trp*) operon** ([\[link\]](#)). If tryptophan is present in the environment, then *E. coli* does not need to synthesize it and the switch controlling the activation of the genes in the *trp* operon is switched off. However, when tryptophan availability is low, the switch controlling the operon is turned on, transcription is initiated, the genes are expressed, and tryptophan is synthesized.



The five genes that are needed to synthesize tryptophan in *E. coli* are located next to each other in the *trp* operon. When tryptophan is plentiful, two tryptophan molecules bind the repressor protein at the operator sequence. This physically blocks the RNA polymerase from transcribing the tryptophan genes. When tryptophan is absent, the repressor protein does not bind to the operator and the genes are transcribed.

A DNA sequence that codes for proteins is referred to as the coding region. The five coding regions for the tryptophan biosynthesis enzymes are arranged sequentially on the chromosome in the operon. Just before the coding region is the **transcriptional start site**. This is the region of DNA to which RNA polymerase binds to initiate transcription. The promoter sequence is upstream of the transcriptional start site; each operon has a sequence within or near the promoter to which proteins (activators or repressors) can bind and regulate transcription.

A DNA sequence called the operator sequence is encoded between the promoter region and the first *trp* coding gene. This **operator** contains the

DNA code to which the repressor protein can bind. When tryptophan is present in the cell, two tryptophan molecules bind to the *trp* repressor, which changes shape to bind to the *trp* operator. Binding of the tryptophan–repressor complex at the operator physically prevents the RNA polymerase from binding, and transcribing the downstream genes.

When tryptophan is not present in the cell, the repressor by itself does not bind to the operator; therefore, the operon is active and tryptophan is synthesized. Because the repressor protein actively binds to the operator to keep the genes turned off, the *trp* operon is negatively regulated and the proteins that bind to the operator to silence *trp* expression are **negative regulators**.

**Note:**

Link to Learning



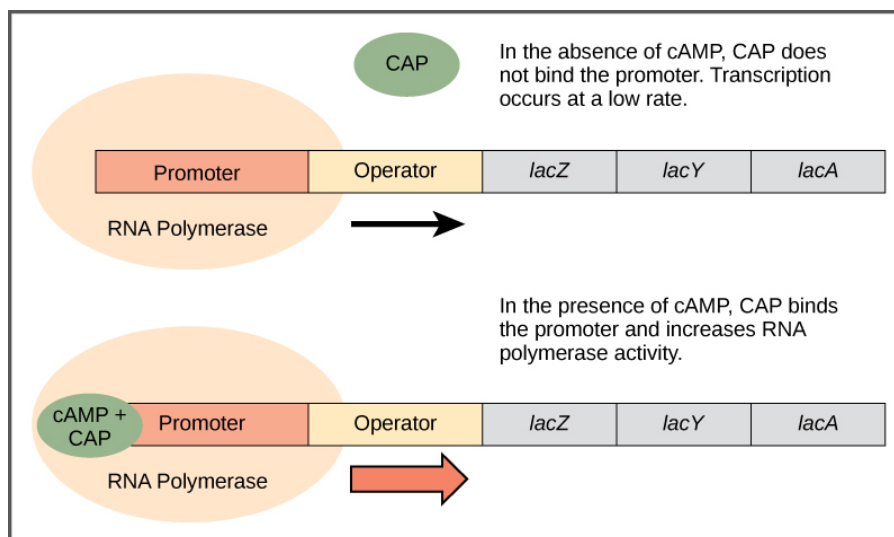
Watch this video to learn more about the *trp* operon.

[https://www.openstaxcollege.org/l/trp\\_operon](https://www.openstaxcollege.org/l/trp_operon)

## Catabolite Activator Protein (CAP): An Activator Regulator

Just as the *trp* operon is negatively regulated by tryptophan molecules, there are proteins that bind to the operator sequences that act as a **positive regulator** to turn genes on and activate them. For example, when glucose is scarce, *E. coli* bacteria can turn to other sugar sources for fuel. To do this, new genes to process these alternate genes must be transcribed. When glucose levels drop, cyclic AMP (cAMP) begins to accumulate in the cell.

The cAMP molecule is a signaling molecule that is involved in glucose and energy metabolism in *E. coli*. When glucose levels decline in the cell, accumulating cAMP binds to the positive regulator **catabolite activator protein (CAP)**, a protein that binds to the promoters of operons that control the processing of alternative sugars. When cAMP binds to CAP, the complex binds to the promoter region of the genes that are needed to use the alternate sugar sources ([\[link\]](#)). In these operons, a CAP binding site is located upstream of the RNA polymerase binding site in the promoter. This increases the binding ability of RNA polymerase to the promoter region and the transcription of the genes.



When glucose levels fall, *E. coli* may use other sugars for fuel but must transcribe new genes to do so. As glucose supplies become limited, cAMP levels increase. This cAMP binds to the CAP protein, a positive regulator that binds to an operator region upstream of the genes required to use other sugar sources.

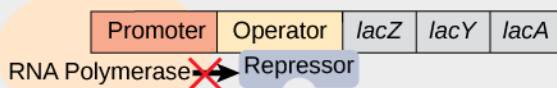
## The *lac* Operon: An Inducer Operon

The third type of gene regulation in prokaryotic cells occurs through **inducible operons**, which have proteins that bind to activate or repress transcription depending on the local environment and the needs of the cell. The *lac* operon is a typical inducible operon. As mentioned previously, *E. coli* is able to use other sugars as energy sources when glucose concentrations are low. To do so, the cAMP–CAP protein complex serves as a positive regulator to induce transcription. One such sugar source is lactose. The ***lac* operon** encodes the genes necessary to acquire and process the lactose from the local environment. CAP binds to the operator sequence upstream of the promoter that initiates transcription of the *lac* operon. However, for the *lac* operon to be activated, two conditions must be met. First, the level of glucose must be very low or non-existent. Second, lactose must be present. Only when glucose is absent and lactose is present will the *lac* operon be transcribed ([\[link\]](#)). This makes sense for the cell, because it would be energetically wasteful to create the proteins to process lactose if glucose was plentiful or lactose was not available.

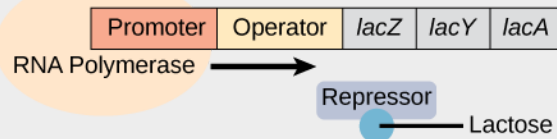
**Note:**

Art Connection

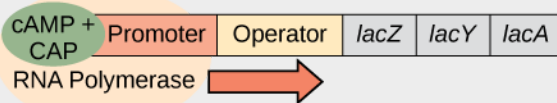
In the absence of lactose, the lac repressor binds the operator, and transcription is blocked.



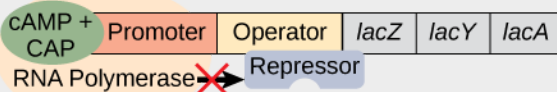
In the presence of lactose, the lac repressor is released from the operator, and transcription proceeds at a slow rate.



cAMP-CAP complex stimulates RNA Polymerase activity and increases RNA synthesis.



However, even in the presence of cAMP-CAP complex, RNA synthesis is blocked when repressor is bound to the operator.



Transcription of the *lac* operon is carefully regulated so that its expression only occurs when glucose is limited and lactose is present to serve as an alternative fuel source.

In *E. coli*, the *trp* operon is on by default, while the *lac* operon is off. Why do you think this is the case?

If glucose is absent, then CAP can bind to the operator sequence to activate transcription. If lactose is absent, then the repressor binds to the operator to prevent transcription. If either of these requirements is met, then transcription remains off. Only when both conditions are satisfied is the *lac* operon transcribed ([link](#)).

Signals that Induce or Repress Transcription of the <i>lac</i> Operon				
Glucose	CAP binds	Lactose	Repressor binds	Transcription
+	-	-	+	No
+	-	+	-	Some
-	+	-	+	No
-	+	+	-	Yes

**Note:**

Link to Learning



Watch an animated tutorial about the workings of *lac* operon here.

[https://www.openstaxcollege.org/l/lac\\_operon](https://www.openstaxcollege.org/l/lac_operon)

## Section Summary

The regulation of gene expression in prokaryotic cells occurs at the transcriptional level. There are three ways to control the transcription of an operon: repressive control, activator control, and inducible control.

Repressive control, typified by the *trp* operon, uses proteins bound to the operator sequence to physically prevent the binding of RNA polymerase and the activation of transcription. Therefore, if tryptophan is not needed, the repressor is bound to the operator and transcription remains off.

Activator control, typified by the action of CAP, increases the binding ability of RNA polymerase to the promoter when CAP is bound. In this case, low levels of glucose result in the binding of cAMP to CAP. CAP then binds the promoter, which allows RNA polymerase to bind to the promoter better. In the last example—the *lac* operon—two conditions must be met to initiate transcription. Glucose must not be present, and lactose must be available for the *lac* operon to be transcribed. If glucose is absent, CAP binds to the operator. If lactose is present, the repressor protein does not bind to its operator. Only when both conditions are met will RNA polymerase bind to the promoter to induce transcription.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) In *E. coli*, the *trp* operon is on by default, while the *lac* operon is off. Why do you think that this is the case?

---

#### Solution:

[\[link\]](#) Tryptophan is an amino acid essential for making proteins, so the cell always needs to have some on hand. However, if plenty of tryptophan is present, it is wasteful to make more, and the expression of the *trp* receptor is repressed. Lactose, a sugar found in milk, is not always available. It makes no sense to make the enzymes necessary to digest an energy source that is not available, so the *lac* operon is only turned on when lactose is present.



## Review Questions

### Exercise:

#### Problem:

If glucose is absent, but so is lactose, the *lac* operon will be \_\_\_\_\_.

- a. activated
- b. repressed
- c. activated, but only partially
- d. mutated

---

#### Solution:

B

### Exercise:

#### Problem:

Prokaryotic cells lack a nucleus. Therefore, the genes in prokaryotic cells are:

- a. all expressed, all of the time
- b. transcribed and translated almost simultaneously
- c. transcriptionally controlled because translation begins before transcription ends
- d. b and c are both true

---

#### Solution:

D

## Free Response

**Exercise:****Problem:**

Describe how transcription in prokaryotic cells can be altered by external stimulation such as excess lactose in the environment.

---

**Solution:**

Environmental stimuli can increase or induce transcription in prokaryotic cells. In this example, lactose in the environment will induce the transcription of the *lac* operon, but only if glucose is not available in the environment.

**Exercise:****Problem:**

What is the difference between a repressible and an inducible operon?

---

**Solution:**

A repressible operon uses a protein bound to the promoter region of a gene to keep the gene repressed or silent. This repressor must be actively removed in order to transcribe the gene. An inducible operon is either activated or repressed depending on the needs of the cell and what is available in the local environment.

**Glossary**

activator

protein that binds to prokaryotic operators to increase transcription

catabolite activator protein (CAP)

protein that complexes with cAMP to bind to the promoter sequences of operons that control sugar processing when glucose is not available

inducible operon

operon that can be activated or repressed depending on cellular needs and the surrounding environment

lac operon

operon in prokaryotic cells that encodes genes required for processing and intake of lactose

negative regulator

protein that prevents transcription

operator

region of DNA outside of the promoter region that binds activators or repressors that control gene expression in prokaryotic cells

operon

collection of genes involved in a pathway that are transcribed together as a single mRNA in prokaryotic cells

positive regulator

protein that increases transcription

repressor

protein that binds to the operator of prokaryotic genes to prevent transcription

transcriptional start site

site at which transcription begins

trp operon

series of genes necessary to synthesize tryptophan in prokaryotic cells

tryptophan

amino acid that can be synthesized by prokaryotic cells when necessary

## Eukaryotic Epigenetic Gene Regulation

By the end of this section, you will be able to:

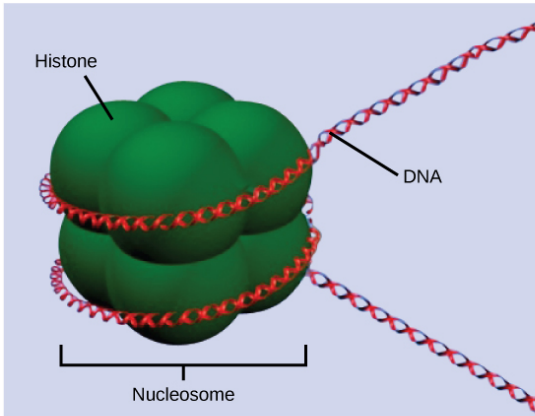
- Explain the process of epigenetic regulation
- Describe how access to DNA is controlled by histone modification

Eukaryotic gene expression is more complex than prokaryotic gene expression because the processes of transcription and translation are physically separated. Unlike prokaryotic cells, eukaryotic cells can regulate gene expression at many different levels. Eukaryotic gene expression begins with control of access to the DNA. This form of regulation, called epigenetic regulation, occurs even before transcription is initiated.

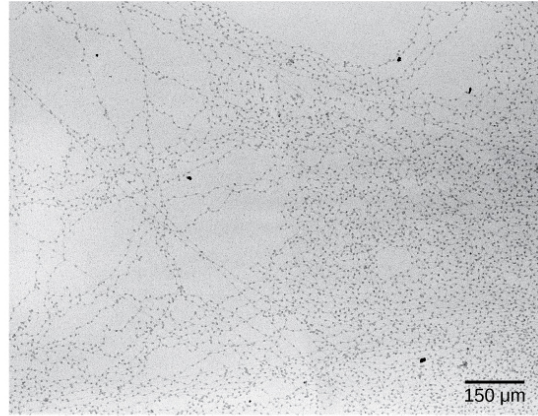
### **Epigenetic Control: Regulating Access to Genes within the Chromosome**

The human genome encodes over 20,000 genes; each of the 23 pairs of human chromosomes encodes thousands of genes. The DNA in the nucleus is precisely wound, folded, and compacted into chromosomes so that it will fit into the nucleus. It is also organized so that specific segments can be accessed as needed by a specific cell type.

The first level of organization, or packing, is the winding of DNA strands around histone proteins. Histones package and order DNA into structural units called nucleosome complexes, which can control the access of proteins to the DNA regions ([\[link\]](#)**a**). Under the electron microscope, this winding of DNA around histone proteins to form nucleosomes looks like small beads on a string ([\[link\]](#)**b**). These beads (histone proteins) can move along the string (DNA) and change the structure of the molecule.



(a)



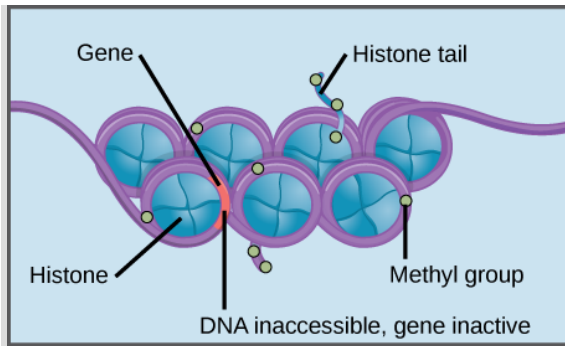
(b)

DNA is folded around histone proteins to create (a) nucleosome complexes. These nucleosomes control the access of proteins to the underlying DNA. When viewed through an electron microscope (b), the nucleosomes look like beads on a string. (credit “micrograph”: modification of work by Chris Woodcock)

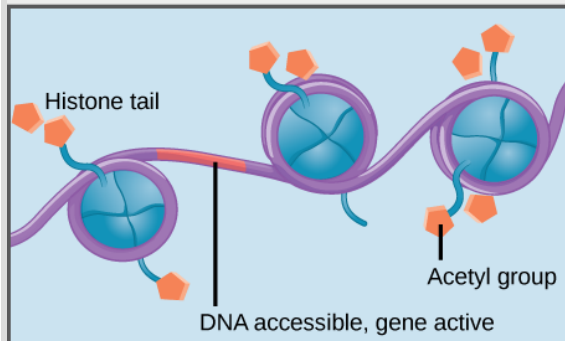
If DNA encoding a specific gene is to be transcribed into RNA, the nucleosomes surrounding that region of DNA can slide down the DNA to open that specific chromosomal region and allow for the transcriptional machinery (RNA polymerase) to initiate transcription ([\[link\]](#)). Nucleosomes can move to open the chromosome structure to expose a segment of DNA, but do so in a very controlled manner.

### **Note:**

Art Connection



Methylation of DNA and histones causes nucleosomes to pack tightly together. Transcription factors cannot bind the DNA, and genes are not expressed.



Histone acetylation results in loose packing of nucleosomes. Transcription factors can bind the DNA and genes are expressed.

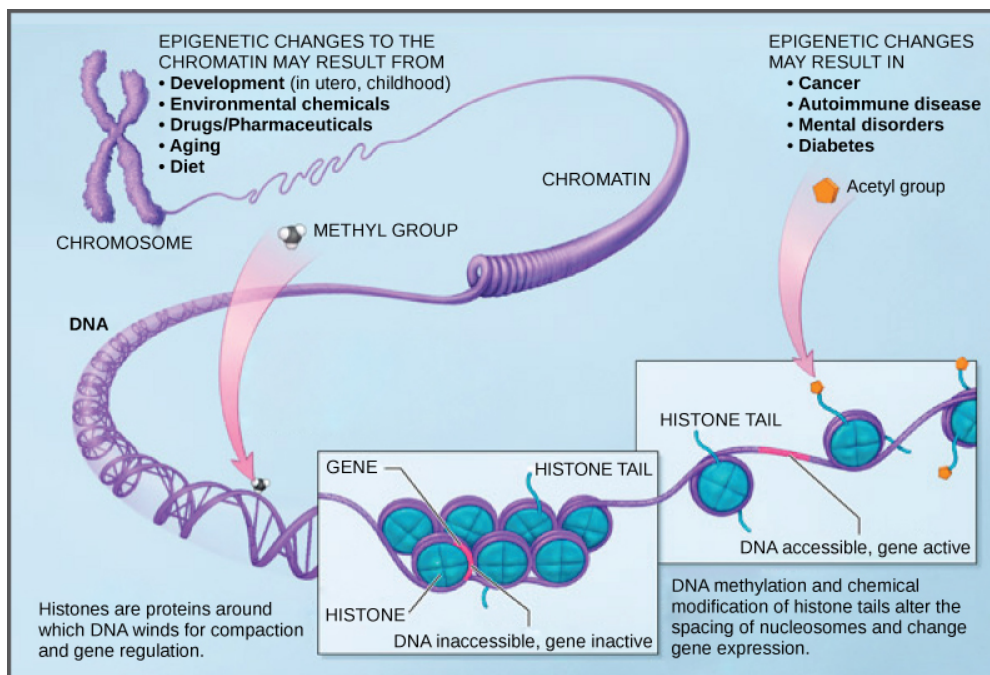
Nucleosomes can slide along DNA. When nucleosomes are spaced closely together (top), transcription factors cannot bind and gene expression is turned off. When the nucleosomes are spaced far apart (bottom), the DNA is exposed. Transcription factors can bind, allowing gene expression to occur. Modifications to the histones and DNA affect nucleosome spacing.

In females, one of the two X chromosomes is inactivated during embryonic development because of epigenetic changes to the chromatin. What impact do you think these changes would have on nucleosome packing?

How the histone proteins move is dependent on signals found on both the histone proteins and on the DNA. These signals are tags added to histone proteins and DNA that tell the histones if a chromosomal region should be open or closed ([link](#) depicts modifications to histone proteins and DNA). These tags are not permanent, but may be added or removed as needed.

They are chemical modifications (phosphate, methyl, or acetyl groups) that are attached to specific amino acids in the protein or to the nucleotides of the DNA. The tags do not alter the DNA base sequence, but they do alter how tightly wound the DNA is around the histone proteins. DNA is a negatively charged molecule; therefore, changes in the charge of the histone will change how tightly wound the DNA molecule will be. When unmodified, the histone proteins have a large positive charge; by adding chemical modifications like acetyl groups, the charge becomes less positive.

The DNA molecule itself can also be modified. This occurs within very specific regions called CpG islands. These are stretches with a high frequency of cytosine and guanine dinucleotide DNA pairs (CG) found in the promoter regions of genes. When this configuration exists, the cytosine member of the pair can be methylated (a methyl group is added). This modification changes how the DNA interacts with proteins, including the histone proteins that control access to the region. Highly methylated (hypermethylated) DNA regions with deacetylated histones are tightly coiled and transcriptionally inactive.



Histone proteins and DNA nucleotides can be modified

chemically. Modifications affect nucleosome spacing and gene expression. (credit: modification of work by NIH)

This type of gene regulation is called epigenetic regulation. Epigenetic means “around genetics.” The changes that occur to the histone proteins and DNA do not alter the nucleotide sequence and are not permanent. Instead, these changes are temporary (although they often persist through multiple rounds of cell division) and alter the chromosomal structure (open or closed) as needed. A gene can be turned on or off depending upon the location and modifications to the histone proteins and DNA. If a gene is to be transcribed, the histone proteins and DNA are modified surrounding the chromosomal region encoding that gene. This opens the chromosomal region to allow access for RNA polymerase and other proteins, called **transcription factors**, to bind to the promoter region, located just upstream of the gene, and initiate transcription. If a gene is to remain turned off, or silenced, the histone proteins and DNA have different modifications that signal a closed chromosomal configuration. In this closed configuration, the RNA polymerase and transcription factors do not have access to the DNA and transcription cannot occur ([link](#)).

**Note:**

Link to Learning



View this video that describes how epigenetic regulation controls gene expression.

[https://www.openstaxcollege.org/l/epigenetic\\_reg](https://www.openstaxcollege.org/l/epigenetic_reg)



## Section Summary

In eukaryotic cells, the first stage of gene expression control occurs at the epigenetic level. Epigenetic mechanisms control access to the chromosomal region to allow genes to be turned on or off. These mechanisms control how DNA is packed into the nucleus by regulating how tightly the DNA is wound around histone proteins. The addition or removal of chemical modifications (or flags) to histone proteins or DNA signals to the cell to open or close a chromosomal region. Therefore, eukaryotic cells can control whether a gene is expressed by controlling accessibility to transcription factors and the binding of RNA polymerase to initiate transcription.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) In females, one of the two X chromosomes is inactivated during embryonic development because of epigenetic changes to the chromatin. What impact do you think these changes would have on nucleosome packing?

---

#### Solution:

[\[link\]](#) The nucleosomes would pack more tightly together.

## Review Questions

### Exercise:

**Problem:** What are epigenetic modifications?

- a. the addition of reversible changes to histone proteins and DNA
- b. the removal of nucleosomes from the DNA
- c. the addition of more nucleosomes to the DNA
- d. mutation of the DNA sequence

---

**Solution:**

A

**Exercise:**

**Problem:** Which of the following are true of epigenetic changes?

- a. allow DNA to be transcribed
- b. move histones to open or close a chromosomal region
- c. are temporary
- d. all of the above

---

**Solution:**

D

## Free Response

**Exercise:**

**Problem:**

In cancer cells, alteration to epigenetic modifications turns off genes that are normally expressed. Hypothetically, how could you reverse this process to turn these genes back on?

---

**Solution:**

You can create medications that reverse the epigenetic processes (to add histone acetylation marks or to remove DNA methylation) and create an open chromosomal configuration.

## Glossary

transcription factor

protein that binds to the DNA at the promoter or enhancer region and that influences transcription of a gene

## Eukaryotic Transcription Gene Regulation

By the end of this section, you will be able to:

- Discuss the role of transcription factors in gene regulation
- Explain how enhancers and repressors regulate gene expression

Like prokaryotic cells, the transcription of genes in eukaryotes requires the actions of an RNA polymerase to bind to a sequence upstream of a gene to initiate transcription. However, unlike prokaryotic cells, the eukaryotic RNA polymerase requires other proteins, or transcription factors, to facilitate transcription initiation. Transcription factors are proteins that bind to the promoter sequence and other regulatory sequences to control the transcription of the target gene. RNA polymerase by itself cannot initiate transcription in eukaryotic cells. Transcription factors must bind to the promoter region first and recruit RNA polymerase to the site for transcription to be established.

### **Note:**

Link to Learning



View the process of transcription—the making of RNA from a DNA template.

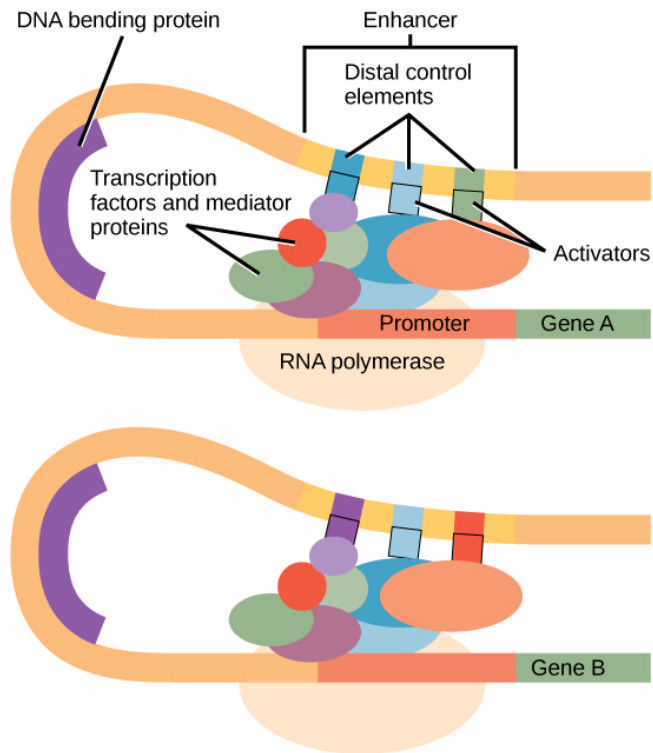
[https://www.openstaxcollege.org/l/transcript\\_RNA](https://www.openstaxcollege.org/l/transcript_RNA)

## The Promoter and the Transcription Machinery

Genes are organized to make the control of gene expression easier. The promoter region is immediately upstream of the coding sequence. This

region can be short (only a few nucleotides in length) or quite long (hundreds of nucleotides long). The longer the promoter, the more available space for proteins to bind. This also adds more control to the transcription process. The length of the promoter is gene-specific and can differ dramatically between genes. Consequently, the level of control of gene expression can also differ quite dramatically between genes. The purpose of the promoter is to bind transcription factors that control the initiation of transcription.

Within the promoter region, just upstream of the transcriptional start site, resides the TATA box. This box is simply a repeat of thymine and adenine dinucleotides (literally, TATA repeats). RNA polymerase binds to the transcription initiation complex, allowing transcription to occur. To initiate transcription, a transcription factor (TFIID) is the first to bind to the TATA box. Binding of TFIID recruits other transcription factors, including TFIIB, TFIIE, TFIIIF, and TFIIH to the TATA box. Once this complex is assembled, RNA polymerase can bind to its upstream sequence. When bound along with the transcription factors, RNA polymerase is phosphorylated. This releases part of the protein from the DNA to activate the transcription initiation complex and places RNA polymerase in the correct orientation to begin transcription; DNA-bending protein brings the enhancer, which can be quite a distance from the gene, in contact with transcription factors and mediator proteins ([\[link\]](#)).



An enhancer is a DNA sequence that promotes transcription. Each enhancer is made up of short DNA sequences called distal control elements. Activators bound to the distal control elements interact with mediator proteins and transcription factors. Two different genes may have the same promoter but different distal control elements, enabling differential gene expression.

In addition to the general transcription factors, other transcription factors can bind to the promoter to regulate gene transcription. These transcription factors bind to the promoters of a specific set of genes. They are not general transcription factors that bind to every promoter complex, but are recruited to a specific sequence on the promoter of a specific gene. There are

hundreds of transcription factors in a cell that each bind specifically to a particular DNA sequence motif. When transcription factors bind to the promoter just upstream of the encoded gene, it is referred to as a **cis-acting element**, because it is on the same chromosome just next to the gene. The region that a particular transcription factor binds to is called the **transcription factor binding site**. Transcription factors respond to environmental stimuli that cause the proteins to find their binding sites and initiate transcription of the gene that is needed.

## Enhancers and Transcription

In some eukaryotic genes, there are regions that help increase or enhance transcription. These regions, called **enhancers**, are not necessarily close to the genes they enhance. They can be located upstream of a gene, within the coding region of the gene, downstream of a gene, or may be thousands of nucleotides away.

Enhancer regions are binding sequences, or sites, for transcription factors. When a DNA-bending protein binds, the shape of the DNA changes ([link](#)). This shape change allows for the interaction of the activators bound to the enhancers with the transcription factors bound to the promoter region and the RNA polymerase. Whereas DNA is generally depicted as a straight line in two dimensions, it is actually a three-dimensional object. Therefore, a nucleotide sequence thousands of nucleotides away can fold over and interact with a specific promoter.

## Turning Genes Off: Transcriptional Repressors

Like prokaryotic cells, eukaryotic cells also have mechanisms to prevent transcription. Transcriptional repressors can bind to promoter or enhancer regions and block transcription. Like the transcriptional activators, repressors respond to external stimuli to prevent the binding of activating transcription factors.

## Section Summary

To start transcription, general transcription factors, such as TFIID, TFIIF, and others, must first bind to the TATA box and recruit RNA polymerase to that location. The binding of additional regulatory transcription factors to *cis*-acting elements will either increase or prevent transcription. In addition to promoter sequences, enhancer regions help augment transcription. Enhancers can be upstream, downstream, within a gene itself, or on other chromosomes. Transcription factors bind to enhancer regions to increase or prevent transcription.

## Review Questions

### Exercise:

#### Problem:

The binding of \_\_\_\_\_ is required for transcription to start.

- a. a protein
- b. DNA polymerase
- c. RNA polymerase
- d. a transcription factor

---

#### Solution:

C

### Exercise:

#### Problem:

What will result from the binding of a transcription factor to an enhancer region?

- a. decreased transcription of an adjacent gene
  - b. increased transcription of a distant gene
  - c. alteration of the translation of an adjacent gene
  - d. initiation of the recruitment of RNA polymerase
-



**Solution:**

B

**Free Response****Exercise:****Problem:**

A mutation within the promoter region can alter transcription of a gene. Describe how this can happen.

---

**Solution:**

A mutation in the promoter region can change the binding site for a transcription factor that normally binds to increase transcription. The mutation could either decrease the ability of the transcription factor to bind, thereby decreasing transcription, or it can increase the ability of the transcription factor to bind, thus increasing transcription.

**Exercise:****Problem:**

What could happen if a cell had too much of an activating transcription factor present?

---

**Solution:**

If too much of an activating transcription factor were present, then transcription would be increased in the cell. This could lead to dramatic alterations in cell function.

**Glossary**

*cis*-acting element

transcription factor binding sites within the promoter that regulate the transcription of a gene adjacent to it

enhancer

segment of DNA that is upstream, downstream, perhaps thousands of nucleotides away, or on another chromosome that influence the transcription of a specific gene

*trans*-acting element

transcription factor binding site found outside the promoter or on another chromosome that influences the transcription of a particular gene

transcription factor binding site

sequence of DNA to which a transcription factor binds

## Eukaryotic Post-transcriptional Gene Regulation

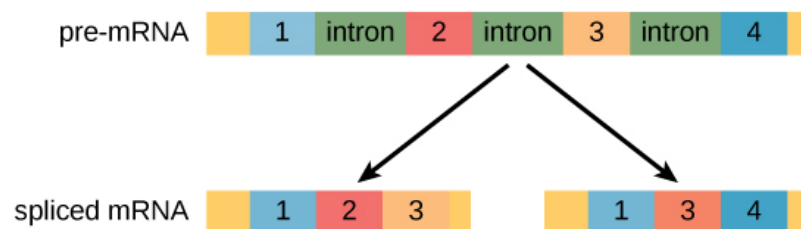
By the end of this section, you will be able to:

- Understand RNA splicing and explain its role in regulating gene expression
- Describe the importance of RNA stability in gene regulation

RNA is transcribed, but must be processed into a mature form before translation can begin. This processing after an RNA molecule has been transcribed, but before it is translated into a protein, is called post-transcriptional modification. As with the epigenetic and transcriptional stages of processing, this post-transcriptional step can also be regulated to control gene expression in the cell. If the RNA is not processed, shuttled, or translated, then no protein will be synthesized.

### RNA splicing, the first stage of post-transcriptional control

In eukaryotic cells, the RNA transcript often contains regions, called introns, that are removed prior to translation. The regions of RNA that code for protein are called exons ([\[link\]](#)). After an RNA molecule has been transcribed, but prior to its departure from the nucleus to be translated, the RNA is processed and the introns are removed by splicing.

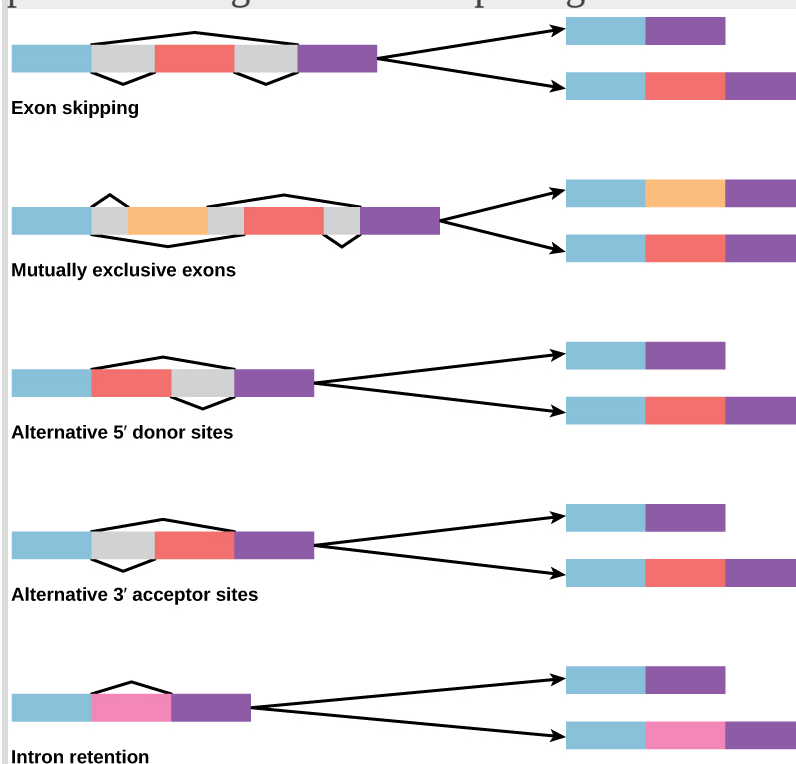


Pre-mRNA can be alternatively spliced to create different proteins.

---

**Note:****Evolution Connection****Alternative RNA Splicing**

In the 1970s, genes were first observed that exhibited alternative RNA splicing. Alternative RNA splicing is a mechanism that allows different protein products to be produced from one gene when different combinations of introns, and sometimes exons, are removed from the transcript ([\[link\]](#)). This alternative splicing can be haphazard, but more often it is controlled and acts as a mechanism of gene regulation, with the frequency of different splicing alternatives controlled by the cell as a way to control the production of different protein products in different cells or at different stages of development. Alternative splicing is now understood to be a common mechanism of gene regulation in eukaryotes; according to one estimate, 70 percent of genes in humans are expressed as multiple proteins through alternative splicing.



There are five basic modes of alternative splicing.

How could alternative splicing evolve? Introns have a beginning and ending recognition sequence; it is easy to imagine the failure of the splicing mechanism to identify the end of an intron and instead find the end of the next intron, thus removing two introns and the intervening exon. In fact, there are mechanisms in place to prevent such intron skipping, but mutations are likely to lead to their failure. Such “mistakes” would more than likely produce a nonfunctional protein. Indeed, the cause of many genetic diseases is alternative splicing rather than mutations in a sequence. However, alternative splicing would create a protein variant without the loss of the original protein, opening up possibilities for adaptation of the new variant to new functions. Gene duplication has played an important role in the evolution of new functions in a similar way by providing genes that may evolve without eliminating the original, functional protein.

**Note:**

Link to Learning



Visualize how mRNA splicing happens by watching the process in action in this video.

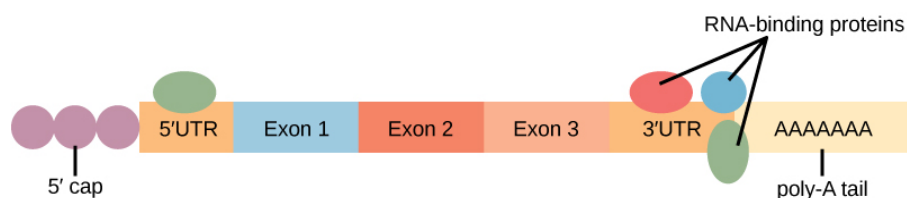
[https://www.openstaxcollege.org/l/mRNA\\_splicing](https://www.openstaxcollege.org/l/mRNA_splicing)

## Control of RNA Stability

Before the mRNA leaves the nucleus, it is given two protective "caps" that prevent the end of the strand from degrading during its journey. The **5' cap**, which is placed on the 5' end of the mRNA, is usually composed of a methylated guanosine triphosphate molecule (GTP). The **poly-A tail**, which

is attached to the 3' end, is usually composed of a series of adenine nucleotides. Once the RNA is transported to the cytoplasm, the length of time that the RNA resides there can be controlled. Each RNA molecule has a defined lifespan and decays at a specific rate. This rate of decay can influence how much protein is in the cell. If the decay rate is increased, the RNA will not exist in the cytoplasm as long, shortening the time for translation to occur. Conversely, if the rate of decay is decreased, the RNA molecule will reside in the cytoplasm longer and more protein can be translated. This rate of decay is referred to as the RNA stability. If the RNA is stable, it will be detected for longer periods of time in the cytoplasm.

Binding of proteins to the RNA can influence its stability. Proteins, called **RNA-binding proteins**, or RBPs, can bind to the regions of the RNA just upstream or downstream of the protein-coding region. These regions in the RNA that are not translated into protein are called the **untranslated regions**, or UTRs. They are not introns (those have been removed in the nucleus). Rather, these are regions that regulate mRNA localization, stability, and protein translation. The region just before the protein-coding region is called the **5' UTR**, whereas the region after the coding region is called the **3' UTR** ([\[link\]](#)). The binding of RBPs to these regions can increase or decrease the stability of an RNA molecule, depending on the specific RBP that binds.



The protein-coding region of mRNA is flanked by 5' and 3' untranslated regions (UTRs). The presence of RNA-binding proteins at the 5' or 3' UTR influences the stability of the RNA molecule.

## RNA Stability and microRNAs

In addition to RBPs that bind to and control (increase or decrease) RNA stability, other elements called microRNAs can bind to the RNA molecule. These **microRNAs**, or miRNAs, are short RNA molecules that are only 21–24 nucleotides in length. The miRNAs are made in the nucleus as longer pre-miRNAs. These pre-miRNAs are chopped into mature miRNAs by a protein called **dicer**. Like transcription factors and RBPs, mature miRNAs recognize a specific sequence and bind to the RNA; however, miRNAs also associate with a ribonucleoprotein complex called the **RNA-induced silencing complex (RISC)**. RISC binds along with the miRNA to degrade the target mRNA. Together, miRNAs and the RISC complex rapidly destroy the RNA molecule.

## Section Summary

Post-transcriptional control can occur at any stage after transcription, including RNA splicing, nuclear shuttling, and RNA stability. Once RNA is transcribed, it must be processed to create a mature RNA that is ready to be translated. This involves the removal of introns that do not code for protein. Spliceosomes bind to the signals that mark the exon/intron border to remove the introns and ligate the exons together. Once this occurs, the RNA is mature and can be translated. RNA is created and spliced in the nucleus, but needs to be transported to the cytoplasm to be translated. RNA is transported to the cytoplasm through the nuclear pore complex. Once the RNA is in the cytoplasm, the length of time it resides there before being degraded, called RNA stability, can also be altered to control the overall amount of protein that is synthesized. The RNA stability can be increased, leading to longer residency time in the cytoplasm, or decreased, leading to shortened time and less protein synthesis. RNA stability is controlled by RNA-binding proteins (RBPs) and microRNAs (miRNAs). These RBPs and miRNAs bind to the 5' UTR or the 3' UTR of the RNA to increase or decrease RNA stability. Depending on the RBP, the stability can be increased or decreased significantly; however, miRNAs always decrease stability and promote decay.

## Review Questions

### Exercise:

#### Problem:

Which of the following are involved in post-transcriptional control?

- a. control of RNA splicing
- b. control of RNA shuttling
- c. control of RNA stability
- d. all of the above

---

#### Solution:

D

### Exercise:

#### Problem:

Binding of an RNA binding protein will \_\_\_\_\_ the stability of the RNA molecule.

- a. increase
- b. decrease
- c. neither increase nor decrease
- d. either increase or decrease

---

#### Solution:

D

## Free Response

### Exercise:



**Problem:**

Describe how RBPs can prevent miRNAs from degrading an RNA molecule.

---

**Solution:**

RNA binding proteins (RBP) bind to the RNA and can either increase or decrease the stability of the RNA. If they increase the stability of the RNA molecule, the RNA will remain intact in the cell for a longer period of time than normal. Since both RBPs and miRNAs bind to the RNA molecule, RBP can potentially bind first to the RNA and prevent the binding of the miRNA that will degrade it.

**Exercise:****Problem:**

How can external stimuli alter post-transcriptional control of gene expression?

---

**Solution:**

External stimuli can modify RNA-binding proteins (i.e., through phosphorylation of proteins) to alter their activity.

**Glossary****3' UTR**

3' untranslated region; region just downstream of the protein-coding region in an RNA molecule that is not translated

**5' cap**

a methylated guanosine triphosphate (GTP) molecule that is attached to the 5' end of a messenger RNA to protect the end from degradation

**5' UTR**

5' untranslated region; region just upstream of the protein-coding region in an RNA molecule that is not translated

dicer

enzyme that chops the pre-miRNA into the mature form of the miRNA

microRNA (miRNA)

small RNA molecules (approximately 21 nucleotides in length) that bind to RNA molecules to degrade them

poly-A tail

a series of adenine nucleotides that are attached to the 3' end of an mRNA to protect the end from degradation

RNA-binding protein (RBP)

protein that binds to the 3' or 5' UTR to increase or decrease the RNA stability

RNA stability

how long an RNA molecule will remain intact in the cytoplasm

untranslated region

segment of the RNA molecule that are not translated into protein. These regions lie before (upstream or 5') and after (downstream or 3') the protein-coding region

RISC

protein complex that binds along with the miRNA to the RNA to degrade it

## Eukaryotic Translational and Post-translational Gene Regulation

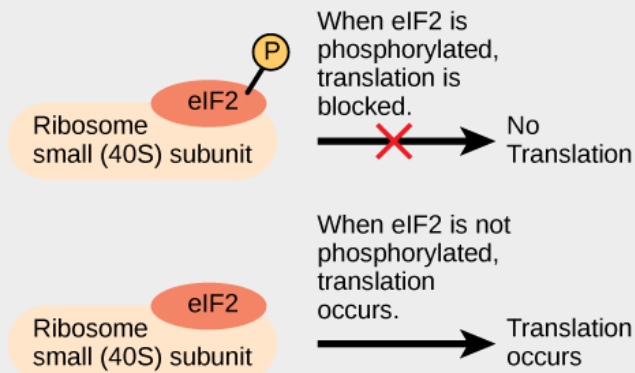
By the end of this section, you will be able to:

- Understand the process of translation and discuss its key factors
- Describe how the initiation complex controls translation
- Explain the different ways in which the post-translational control of gene expression takes place

After the RNA has been transported to the cytoplasm, it is translated into protein. Control of this process is largely dependent on the RNA molecule. As previously discussed, the stability of the RNA will have a large impact on its translation into a protein. As the stability changes, the amount of time that it is available for translation also changes.

### The Initiation Complex and Translation Rate

Like transcription, translation is controlled by proteins that bind and initiate the process. In translation, the complex that assembles to start the process is referred to as the **initiation complex**. The first protein to bind to the RNA to initiate translation is the **eukaryotic initiation factor-2 (eIF-2)**. The eIF-2 protein is active when it binds to the high-energy molecule **guanosine triphosphate (GTP)**. GTP provides the energy to start the reaction by giving up a phosphate and becoming **guanosine diphosphate (GDP)**. The eIF-2 protein bound to GTP binds to the small **40S ribosomal subunit**. When bound, the methionine initiator tRNA associates with the eIF-2/40S ribosome complex, bringing along with it the mRNA to be translated. At this point, when the initiator complex is assembled, the GTP is converted into GDP and energy is released. The phosphate and the eIF-2 protein are released from the complex and the large **60S ribosomal subunit** binds to translate the RNA. The binding of eIF-2 to the RNA is controlled by phosphorylation. If eIF-2 is phosphorylated, it undergoes a conformational change and cannot bind to GTP. Therefore, the initiation complex cannot form properly and translation is impeded ([\[link\]](#)). When eIF-2 remains unphosphorylated, it binds the RNA and actively translates the protein.

**Note:****Art Connection**

Gene expression can be controlled by factors that bind the translation initiation complex.

An increase in phosphorylation levels of eIF-2 has been observed in patients with neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's. What impact do you think this might have on protein synthesis?

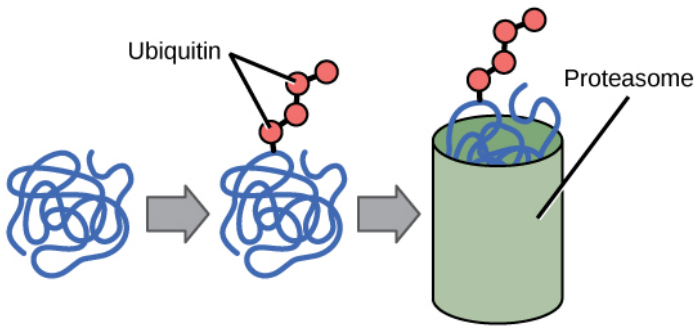
## Chemical Modifications, Protein Activity, and Longevity

Proteins can be chemically modified with the addition of groups including methyl, phosphate, acetyl, and ubiquitin groups. The addition or removal of these groups from proteins regulates their activity or the length of time they exist in the cell. Sometimes these modifications can regulate where a protein is found in the cell—for example, in the nucleus, the cytoplasm, or attached to the plasma membrane.

Chemical modifications occur in response to external stimuli such as stress, the lack of nutrients, heat, or ultraviolet light exposure. These changes can alter epigenetic accessibility, transcription, mRNA stability, or translation—all resulting in changes in expression of various genes. This is an efficient way for the cell to rapidly change the levels of specific proteins in response

to the environment. Because proteins are involved in every stage of gene regulation, the phosphorylation of a protein (depending on the protein that is modified) can alter accessibility to the chromosome, can alter translation (by altering transcription factor binding or function), can change nuclear shuttling (by influencing modifications to the nuclear pore complex), can alter RNA stability (by binding or not binding to the RNA to regulate its stability), can modify translation (increase or decrease), or can change post-translational modifications (add or remove phosphates or other chemical modifications).

The addition of an ubiquitin group to a protein marks that protein for degradation. Ubiquitin acts like a flag indicating that the protein lifespan is complete. These proteins are moved to the **proteasome**, an organelle that functions to remove proteins, to be degraded ([\[link\]](#)). One way to control gene expression, therefore, is to alter the longevity of the protein.



Proteins with ubiquitin tags are marked for degradation within the proteasome.

## Section Summary

Changing the status of the RNA or the protein itself can affect the amount of protein, the function of the protein, or how long it is found in the cell. To translate the protein, a protein initiator complex must assemble on the RNA.

Modifications (such as phosphorylation) of proteins in this complex can prevent proper translation from occurring. Once a protein has been synthesized, it can be modified (phosphorylated, acetylated, methylated, or ubiquitinated). These post-translational modifications can greatly impact the stability, degradation, or function of the protein.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) An increase in phosphorylation levels of eIF-2 has been observed in patients with neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's. What impact do you think this might have on protein synthesis?

---

#### Solution:

[\[link\]](#) Protein synthesis would be inhibited.

## Review Questions

### Exercise:

#### Problem:

Post-translational modifications of proteins can affect which of the following?

- a. protein function
  - b. transcriptional regulation
  - c. chromatin modification
  - d. all of the above
- 

#### Solution:

A

## Free Response

### Exercise:

#### Problem:

Protein modification can alter gene expression in many ways. Describe how phosphorylation of proteins can alter gene expression.

---

#### Solution:

Because proteins are involved in every stage of gene regulation, phosphorylation of a protein (depending on the protein that is modified) can alter accessibility to the chromosome, can alter translation (by altering the transcription factor binding or function), can change nuclear shuttling (by influencing modifications to the nuclear pore complex), can alter RNA stability (by binding or not binding to the RNA to regulate its stability), can modify translation (increase or decrease), or can change post-translational modifications (add or remove phosphates or other chemical modifications).

### Exercise:

#### Problem:

Alternative forms of a protein can be beneficial or harmful to a cell. What do you think would happen if too much of an alternative protein bound to the 3' UTR of an RNA and caused it to degrade?

---

#### Solution:

If the RNA degraded, then less of the protein that the RNA encodes would be translated. This could have dramatic implications for the cell.

### Exercise:

**Problem:**

Changes in epigenetic modifications alter the accessibility and transcription of DNA. Describe how environmental stimuli, such as ultraviolet light exposure, could modify gene expression.

---

**Solution:**

Environmental stimuli, like ultraviolet light exposure, can alter the modifications to the histone proteins or DNA. Such stimuli may change an actively transcribed gene into a silenced gene by removing acetyl groups from histone proteins or by adding methyl groups to DNA.

**Glossary**

eukaryotic initiation factor-2 (eIF-2)

protein that binds first to an mRNA to initiate translation

guanine diphosphate (GDP)

molecule that is left after the energy is used to start translation

guanine triphosphate (GTP)

energy-providing molecule that binds to eIF-2 and is needed for translation

initiation complex

protein complex containing eIF2-2 that starts translation

large 60S ribosomal subunit

second, larger ribosomal subunit that binds to the RNA to translate it into protein

proteasome

organelle that degrades proteins

small 40S ribosomal subunit



ribosomal subunit that binds to the RNA to translate it into protein

## Cancer and Gene Regulation

By the end of this section, you will be able to:

- Describe how changes to gene expression can cause cancer
- Explain how changes to gene expression at different levels can disrupt the cell cycle
- Discuss how understanding regulation of gene expression can lead to better drug design

Cancer is not a single disease but includes many different diseases. In cancer cells, mutations modify cell-cycle control and cells don't stop growing as they normally would. Mutations can also alter the growth rate or the progression of the cell through the cell cycle. One example of a gene modification that alters the growth rate is increased phosphorylation of cyclin B, a protein that controls the progression of a cell through the cell cycle and serves as a cell-cycle checkpoint protein.

For cells to move through each phase of the cell cycle, the cell must pass through checkpoints. This ensures that the cell has properly completed the step and has not encountered any mutation that will alter its function. Many proteins, including cyclin B, control these checkpoints. The phosphorylation of cyclin B, a post-translational event, alters its function. As a result, cells can progress through the cell cycle unimpeded, even if mutations exist in the cell and its growth should be terminated. This post-translational change of cyclin B prevents it from controlling the cell cycle and contributes to the development of cancer.

## **Cancer: Disease of Altered Gene Expression**

Cancer can be described as a disease of altered gene expression. There are many proteins that are turned on or off (gene activation or gene silencing) that dramatically alter the overall activity of the cell. A gene that is not normally expressed in that cell can be switched on and expressed at high levels. This can be the result of gene mutation or changes in gene regulation (epigenetic, transcription, post-transcription, translation, or post-translation).

Changes in epigenetic regulation, transcription, RNA stability, protein translation, and post-translational control can be detected in cancer. While these changes don't occur simultaneously in one cancer, changes at each of these levels can be detected when observing cancer at different sites in different individuals. Therefore, changes in **histone acetylation** (epigenetic modification that leads to gene silencing), activation of transcription factors by phosphorylation, increased RNA stability, increased translational control, and protein modification can all be detected at some point in various cancer cells. Scientists are working to understand the common changes that give rise to certain types of cancer or how a modification might be exploited to destroy a tumor cell.

## **Tumor Suppressor Genes, Oncogenes, and Cancer**

In normal cells, some genes function to prevent excess, inappropriate cell growth. These are tumor suppressor genes, which are active in normal cells to prevent uncontrolled cell growth. There are many tumor suppressor genes in cells. The most studied tumor suppressor gene is p53, which is mutated in over 50 percent of all cancer types. The p53 protein itself functions as a transcription factor. It can bind to sites in the promoters of genes to initiate transcription. Therefore, the mutation of p53 in cancer will dramatically alter the transcriptional activity of its target genes.

### **Note:**

Link to Learning



Watch [this animation](#) to learn more about the use of p53 in fighting cancer.

Proto-oncogenes are positive cell-cycle regulators. When mutated, proto-oncogenes can become oncogenes and cause cancer. Overexpression of the oncogene can lead to uncontrolled cell growth. This is because oncogenes can alter transcriptional activity, stability, or protein translation of another gene that directly or indirectly controls cell growth. An example of an oncogene involved in cancer is a protein called myc. **Myc** is a transcription factor that is aberrantly activated in Burkett's Lymphoma, a cancer of the lymph system. Overexpression of myc transforms normal B cells into cancerous cells that continue to grow uncontrollably. High B-cell numbers can result in tumors that can interfere with normal bodily function. Patients with Burkett's lymphoma can develop tumors on their jaw or in their mouth that interfere with the ability to eat.

## **Cancer and Epigenetic Alterations**

Silencing genes through epigenetic mechanisms is also very common in cancer cells. There are characteristic modifications to histone proteins and DNA that are associated with silenced genes. In cancer cells, the DNA in the promoter region of silenced genes is methylated on cytosine DNA residues in CpG islands. Histone proteins that surround that region lack the acetylation modification that is present when the genes are expressed in normal cells. This combination of DNA methylation and histone deacetylation (epigenetic modifications that lead to gene silencing) is commonly found in cancer. When these modifications occur, the gene present in that chromosomal region is silenced. Increasingly, scientists understand how epigenetic changes are altered in cancer. Because these changes are temporary and can be reversed—for example, by preventing the action of the histone deacetylase protein that removes acetyl groups, or by DNA methyl transferase enzymes that add methyl groups to cytosines in DNA—it is possible to design new drugs and new therapies to take advantage of the reversible nature of these processes. Indeed, many researchers are testing how a silenced gene can be switched back on in a cancer cell to help re-establish normal growth patterns.

Genes involved in the development of many other illnesses, ranging from allergies to inflammation to autism, are thought to be regulated by

epigenetic mechanisms. As our knowledge of how genes are controlled deepens, new ways to treat diseases like cancer will emerge.

## **Cancer and Transcriptional Control**

Alterations in cells that give rise to cancer can affect the transcriptional control of gene expression. Mutations that activate transcription factors, such as increased phosphorylation, can increase the binding of a transcription factor to its binding site in a promoter. This could lead to increased transcriptional activation of that gene that results in modified cell growth. Alternatively, a mutation in the DNA of a promoter or enhancer region can increase the binding ability of a transcription factor. This could also lead to the increased transcription and aberrant gene expression that is seen in cancer cells.

Researchers have been investigating how to control the transcriptional activation of gene expression in cancer. Identifying how a transcription factor binds, or a pathway that activates where a gene can be turned off, has led to new drugs and new ways to treat cancer. In breast cancer, for example, many proteins are overexpressed. This can lead to increased phosphorylation of key transcription factors that increase transcription. One such example is the overexpression of the epidermal growth factor receptor (EGFR) in a subset of breast cancers. The EGFR pathway activates many protein kinases that, in turn, activate many transcription factors that control genes involved in cell growth. New drugs that prevent the activation of EGFR have been developed and are used to treat these cancers.

## **Cancer and Post-transcriptional Control**

Changes in the post-transcriptional control of a gene can also result in cancer. Recently, several groups of researchers have shown that specific cancers have altered expression of miRNAs. Because miRNAs bind to the 3' UTR of RNA molecules to degrade them, overexpression of these miRNAs could be detrimental to normal cellular activity. Too many miRNAs could dramatically decrease the RNA population leading to a decrease in protein expression. Several studies have demonstrated a change in the miRNA population in specific cancer types. It appears that the subset

of miRNAs expressed in breast cancer cells is quite different from the subset expressed in lung cancer cells or even from normal breast cells. This suggests that alterations in miRNA activity can contribute to the growth of breast cancer cells. These types of studies also suggest that if some miRNAs are specifically expressed only in cancer cells, they could be potential drug targets. It would, therefore, be conceivable that new drugs that turn off miRNA expression in cancer could be an effective method to treat cancer.

## **Cancer and Translational/Post-translational Control**

There are many examples of how translational or post-translational modifications of proteins arise in cancer. Modifications are found in cancer cells from the increased translation of a protein to changes in protein phosphorylation to alternative splice variants of a protein. An example of how the expression of an alternative form of a protein can have dramatically different outcomes is seen in colon cancer cells. The c-Flip protein, a protein involved in mediating the cell death pathway, comes in two forms: long (c-FLIPL) and short (c-FLIPS). Both forms appear to be involved in initiating controlled cell death mechanisms in normal cells. However, in colon cancer cells, expression of the long form results in increased cell growth instead of cell death. Clearly, the expression of the wrong protein dramatically alters cell function and contributes to the development of cancer.

## **New Drugs to Combat Cancer: Targeted Therapies**

Scientists are using what is known about the regulation of gene expression in disease states, including cancer, to develop new ways to treat and prevent disease development. Many scientists are designing drugs on the basis of the gene expression patterns within individual tumors. This idea, that therapy and medicines can be tailored to an individual, has given rise to the field of personalized medicine. With an increased understanding of gene regulation and gene function, medicines can be designed to specifically target diseased cells without harming healthy cells. Some new medicines, called targeted therapies, have exploited the overexpression of a specific

protein or the mutation of a gene to develop a new medication to treat disease. One such example is the use of anti-EGF receptor medications to treat the subset of breast cancer tumors that have very high levels of the EGF protein. Undoubtedly, more targeted therapies will be developed as scientists learn more about how gene expression changes can cause cancer.

**Note:**

**Career Connection**

**Clinical Trial Coordinator**

A clinical trial coordinator is the person managing the proceedings of the clinical trial. This job includes coordinating patient schedules and appointments, maintaining detailed notes, building the database to track patients (especially for long-term follow-up studies), ensuring proper documentation has been acquired and accepted, and working with the nurses and doctors to facilitate the trial and publication of the results. A clinical trial coordinator may have a science background, like a nursing degree, or other certification. People who have worked in science labs or in clinical offices are also qualified to become a clinical trial coordinator. These jobs are generally in hospitals; however, some clinics and doctor's offices also conduct clinical trials and may hire a coordinator.

## **Section Summary**

Cancer can be described as a disease of altered gene expression. Changes at every level of eukaryotic gene expression can be detected in some form of cancer at some point in time. In order to understand how changes to gene expression can cause cancer, it is critical to understand how each stage of gene regulation works in normal cells. By understanding the mechanisms of control in normal, non-diseased cells, it will be easier for scientists to understand what goes wrong in disease states including complex ones like cancer.

## **Review Questions**

**Exercise:**

**Problem:** Cancer causing genes are called \_\_\_\_\_.

- a. transformation genes
- b. tumor suppressor genes
- c. oncogenes
- d. mutated genes

---

**Solution:**

C

**Exercise:****Problem:**

Targeted therapies are used in patients with a set gene expression pattern. A targeted therapy that prevents the activation of the estrogen receptor in breast cancer would be beneficial to which type of patient?

- a. patients who express the EGFR receptor in normal cells
- b. patients with a mutation that inactivates the estrogen receptor
- c. patients with lots of the estrogen receptor expressed in their tumor
- d. patients that have no estrogen receptor expressed in their tumor

---

**Solution:**

C

**Free Response****Exercise:**



**Problem:**

New drugs are being developed that decrease DNA methylation and prevent the removal of acetyl groups from histone proteins. Explain how these drugs could affect gene expression to help kill tumor cells.

---

**Solution:**

These drugs will keep the histone proteins and the DNA methylation patterns in the open chromosomal configuration so that transcription is feasible. If a gene is silenced, these drugs could reverse the epigenetic configuration to re-express the gene.

**Exercise:****Problem:**

How can understanding the gene expression pattern in a cancer cell tell you something about that specific form of cancer?

---

**Solution:**

Understanding which genes are expressed in a cancer cell can help diagnose the specific form of cancer. It can also help identify treatment options for that patient. For example, if a breast cancer tumor expresses the EGFR in high numbers, it might respond to specific anti-EGFR therapy. If that receptor is not expressed, it would not respond to that therapy.

**Glossary**

DNA methylation

epigenetic modification that leads to gene silencing; commonly found in cancer cells

histone acetylation

epigenetic modification that leads to gene silencing; commonly found in cancer cells

myc

oncogene that causes cancer in many cancer cells

## Introduction

class="introduction"

In  
genomics,  
the DNA of  
different  
organisms is  
compared,  
enabling  
scientists to  
create maps  
with which  
to navigate  
the DNA of  
different  
organisms.  
(credit  
"map":  
modificatio  
n of photo  
by NASA)



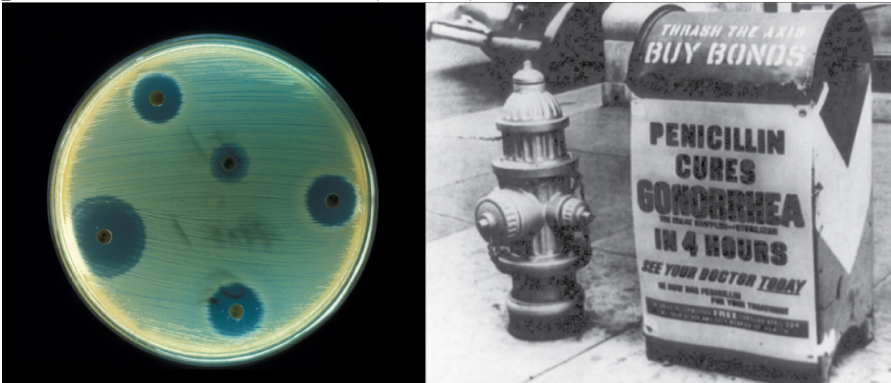
The study of nucleic acids began with the discovery of DNA, progressed to the study of genes and small fragments, and has now exploded to the field of genomics. Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. The advances in genomics have been made possible by DNA sequencing technology. Just as information technology has led to Google maps that enable people to get detailed information about locations around the globe, genomic information is used to create similar maps of the DNA of different organisms. These findings have helped anthropologists to better understand human migration and have aided the field of medicine through the mapping of human genetic diseases. The ways in which genomic information can contribute to scientific understanding are varied and quickly growing.

## Biotechnology

By the end of this section, you will be able to:

- Describe gel electrophoresis
- Explain molecular and reproductive cloning
- Describe uses of biotechnology in medicine and agriculture

**Biotechnology** is the use of biological agents for technological advancement. Biotechnology was used for breeding livestock and crops long before the scientific basis of these techniques was understood. Since the discovery of the structure of DNA in 1953, the field of biotechnology has grown rapidly through both academic research and private companies. The primary applications of this technology are in medicine (production of vaccines and antibiotics) and agriculture (genetic modification of crops, such as to increase yields). Biotechnology also has many industrial applications, such as fermentation, the treatment of oil spills, and the production of biofuels ([\[link\]](#)).



Antibiotics are chemicals produced by fungi, bacteria, and other organisms that have antimicrobial properties. The first antibiotic discovered was penicillin. Antibiotics are now commercially produced and tested for their potential to inhibit bacterial growth. (credit "advertisement": modification of work by NIH; credit "test plate": modification of work by Don Stalons/CDC; scale-bar data from Matt Russell)

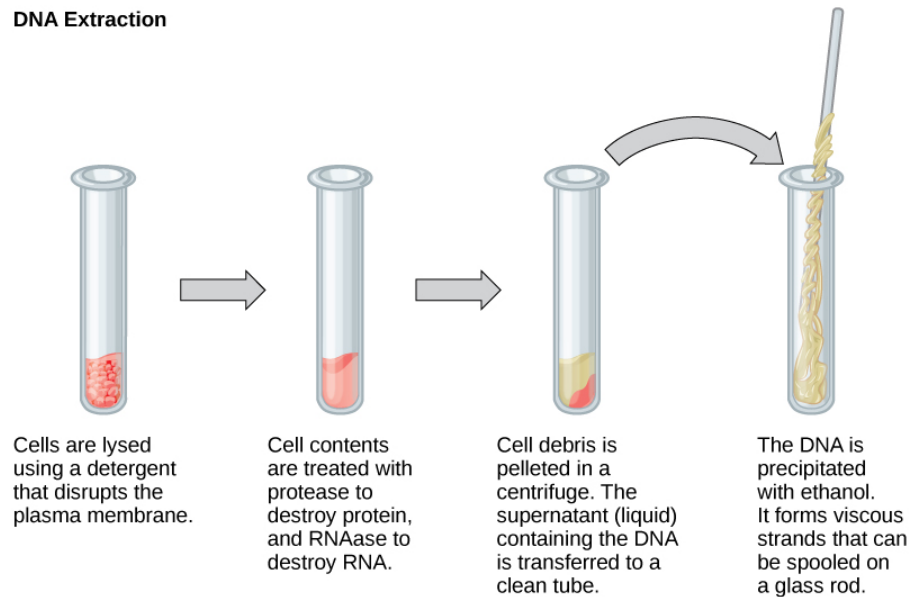
## Basic Techniques to Manipulate Genetic Material (DNA and RNA)

To understand the basic techniques used to work with nucleic acids, remember that nucleic acids are macromolecules made of nucleotides (a sugar, a phosphate, and a nitrogenous base) linked by phosphodiester bonds. The phosphate groups on these molecules each have a net negative charge. An entire set of DNA molecules in the nucleus is called the genome. DNA has two complementary strands linked by hydrogen bonds between the paired bases. The two strands can be separated by exposure to high temperatures (DNA denaturation) and can be reannealed by cooling. The DNA can be replicated by the DNA polymerase enzyme. Unlike DNA, which is located in the nucleus of eukaryotic cells, RNA molecules leave the nucleus. The most common type of RNA that is analyzed is the messenger RNA (mRNA) because it represents the protein-coding genes that are actively expressed. However, RNA molecules present some other challenges to analysis, as they are often less stable than DNA.

### DNA and RNA Extraction

To study or manipulate nucleic acids, the DNA or RNA must first be isolated or extracted from the cells. Various techniques are used to extract different types of DNA ([\[link\]](#)). Most nucleic acid extraction techniques involve steps to break open the cell and use enzymatic reactions to destroy all macromolecules that are not desired (such as degradation of unwanted molecules and separation from the DNA sample). Cells are broken using a **lysis buffer** (a solution which is mostly a detergent); lysis means “to split.” These enzymes break apart lipid molecules in the cell membranes and nuclear membranes. Macromolecules are inactivated using enzymes such as **proteases** that break down proteins, and **ribonucleases** (RNAses) that break down RNA. The DNA is then precipitated using alcohol. Human genomic DNA is usually visible as a gelatinous, white mass. The DNA samples can be stored frozen at  $-80^{\circ}\text{C}$  for several years.

## DNA Extraction



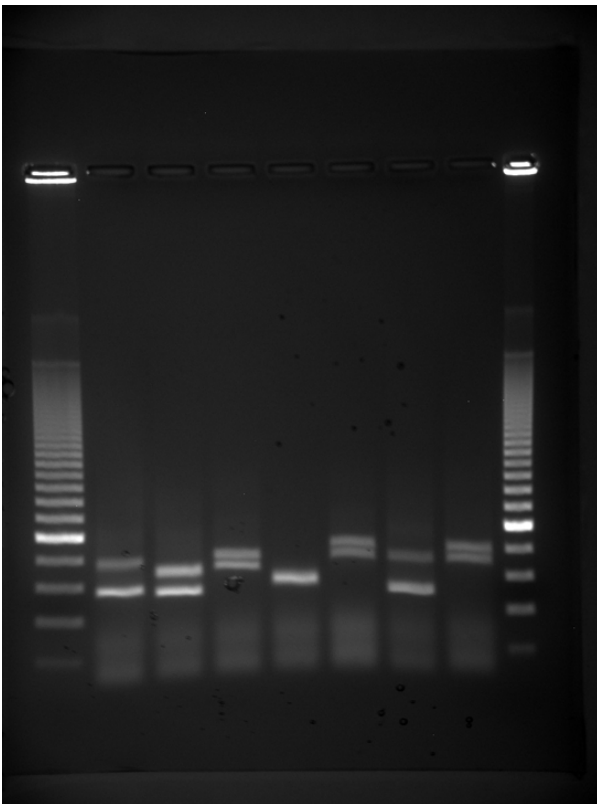
This diagram shows the basic method used for extraction of DNA.

RNA analysis is performed to study gene expression patterns in cells. RNA is naturally very unstable because RNases are commonly present in nature and very difficult to inactivate. Similar to DNA, RNA extraction involves the use of various buffers and enzymes to inactivate macromolecules and preserve the RNA.

## Gel Electrophoresis

Because nucleic acids are negatively charged ions at neutral or basic pH in an aqueous environment, they can be mobilized by an electric field. **Gel electrophoresis** is a technique used to separate molecules on the basis of size, using this charge. The nucleic acids can be separated as whole chromosomes or fragments. The nucleic acids are loaded into a slot near the negative electrode of a semisolid, porous gel matrix and pulled toward the positive electrode at the opposite end of the gel. Smaller molecules move through the pores in the gel faster than larger molecules; this difference in the rate of migration separates the fragments on the basis of size. There are

molecular weight standard samples that can be run alongside the molecules to provide a size comparison. Nucleic acids in a gel matrix can be observed using various fluorescent or colored dyes. Distinct nucleic acid fragments appear as bands at specific distances from the top of the gel (the negative electrode end) on the basis of their size ([\[link\]](#)). A mixture of genomic DNA fragments of varying sizes appear as a long smear, whereas uncut genomic DNA is usually too large to run through the gel and forms a single large band at the top of the gel.



Shown are DNA fragments from seven samples run on a gel, stained with a fluorescent dye, and viewed under UV light. (credit: James Jacob, Tompkins Cortland Community College)

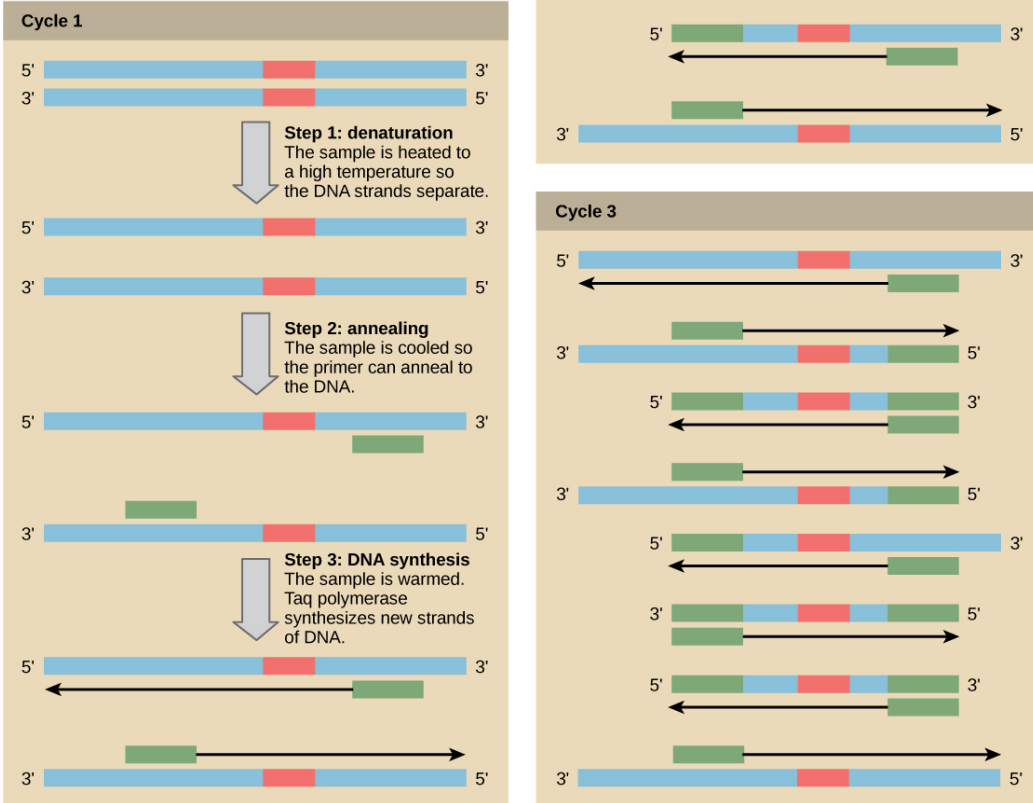


## **Amplification of Nucleic Acid Fragments by Polymerase Chain Reaction**

Although genomic DNA is visible to the naked eye when it is extracted in bulk, DNA analysis often requires focusing on one or more specific regions of the genome. **Polymerase chain reaction (PCR)** is a technique used to amplify specific regions of DNA for further analysis ([\[link\]](#)). PCR is used for many purposes in laboratories, such as the cloning of gene fragments to analyze genetic diseases, identification of contaminant foreign DNA in a sample, and the amplification of DNA for sequencing. More practical applications include the determination of paternity and detection of genetic diseases.

### Polymerase Chain Reaction (PCR)

The PCR cycle consists of three steps—denaturation, annealing, and DNA synthesis—that occur at high, low, and intermediate temperatures, respectively. The cycle is repeated again and again, resulting in a doubling of DNA molecules each time. After several cycles, the vast majority of strands produced are the same length as the distance between the two primers.



Polymerase chain reaction, or PCR, is used to amplify a specific sequence of DNA. Primers—short pieces of DNA complementary to each end of the target sequence—are combined with genomic DNA, Taq polymerase, and deoxynucleotides. Taq polymerase is a DNA polymerase isolated from the thermostable bacterium *Thermus aquaticus* that is able to withstand the high temperatures used in PCR. *Thermus aquaticus* grows in the Lower Geyser Basin of Yellowstone National Park. Reverse transcriptase PCR (RT-PCR) is similar to PCR, but cDNA is made from an RNA template before PCR begins.

DNA fragments can also be amplified from an RNA template in a process called **reverse transcriptase PCR (RT-PCR)**. The first step is to recreate the original DNA template strand (called cDNA) by applying DNA nucleotides to the mRNA. This process is called reverse transcription. This requires the presence of an enzyme called reverse transcriptase. After the cDNA is made, regular PCR can be used to amplify it.

**Note:**

Link to Learning

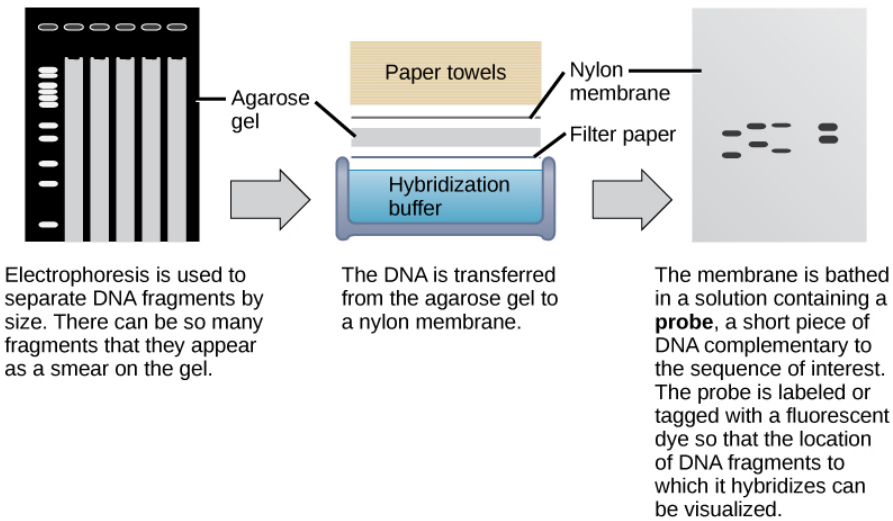


Deepen your understanding of the polymerase chain reaction by clicking through [this interactive exercise](#).

## Hybridization, Southern Blotting, and Northern Blotting

Nucleic acid samples, such as fragmented genomic DNA and RNA extracts, can be probed for the presence of certain sequences. Short DNA fragments called **probes** are designed and labeled with radioactive or fluorescent dyes to aid detection. Gel electrophoresis separates the nucleic acid fragments according to their size. The fragments in the gel are then transferred onto a nylon membrane in a procedure called **blotting** ([link](#)). The nucleic acid fragments that are bound to the surface of the membrane can then be probed with specific radioactively or fluorescently labeled probe sequences. When DNA is transferred to a nylon membrane, the technique is called **Southern blotting**, and when RNA is transferred to a nylon membrane, it is called **northern blotting**. Southern blots are used to detect the presence of certain DNA sequences in a given genome, and northern blots are used to detect gene expression.

### Southern Blotting



Southern blotting is used to find a particular sequence in a sample of DNA. DNA fragments are separated on a gel, transferred to a nylon membrane, and incubated with a DNA probe complementary to the sequence of interest. Northern blotting is similar to Southern blotting, but RNA is run on the gel instead of DNA. In western blotting, proteins are run on a gel and detected using antibodies.

## Molecular Cloning

In general, the word “cloning” means the creation of a perfect replica; however, in biology, the re-creation of a whole organism is referred to as “reproductive cloning.” Long before attempts were made to clone an entire organism, researchers learned how to reproduce desired regions or fragments of the genome, a process that is referred to as molecular cloning.

Cloning small fragments of the genome allows for the manipulation and study of specific genes (and their protein products), or noncoding regions in isolation. A plasmid (also called a vector) is a small circular DNA molecule that replicates independently of the chromosomal DNA. In cloning, the

plasmid molecules can be used to provide a "folder" in which to insert a desired DNA fragment. Plasmids are usually introduced into a bacterial host for proliferation. In the bacterial context, the fragment of DNA from the human genome (or the genome of another organism that is being studied) is referred to as **foreign DNA**, or a transgene, to differentiate it from the DNA of the bacterium, which is called the **host DNA**.

Plasmids occur naturally in bacterial populations (such as *Escherichia coli*) and have genes that can contribute favorable traits to the organism, such as **antibiotic resistance** (the ability to be unaffected by antibiotics). Plasmids have been repurposed and engineered as vectors for molecular cloning and the large-scale production of important reagents, such as insulin and human growth hormone. An important feature of plasmid vectors is the ease with which a foreign DNA fragment can be introduced via the **multiple cloning site (MCS)**. The MCS is a short DNA sequence containing multiple sites that can be cut with different commonly available restriction endonucleases. **Restriction endonucleases** recognize specific DNA sequences and cut them in a predictable manner; they are naturally produced by bacteria as a defense mechanism against foreign DNA. Many restriction endonucleases make staggered cuts in the two strands of DNA, such that the cut ends have a 2- or 4-base single-stranded overhang. Because these overhangs are capable of annealing with complementary overhangs, these are called "sticky ends." Addition of an enzyme called DNA ligase permanently joins the DNA fragments via phosphodiester bonds. In this way, any DNA fragment generated by restriction endonuclease cleavage can be spliced between the two ends of a plasmid DNA that has been cut with the same restriction endonuclease ([\[link\]](#)).

## **Recombinant DNA Molecules**

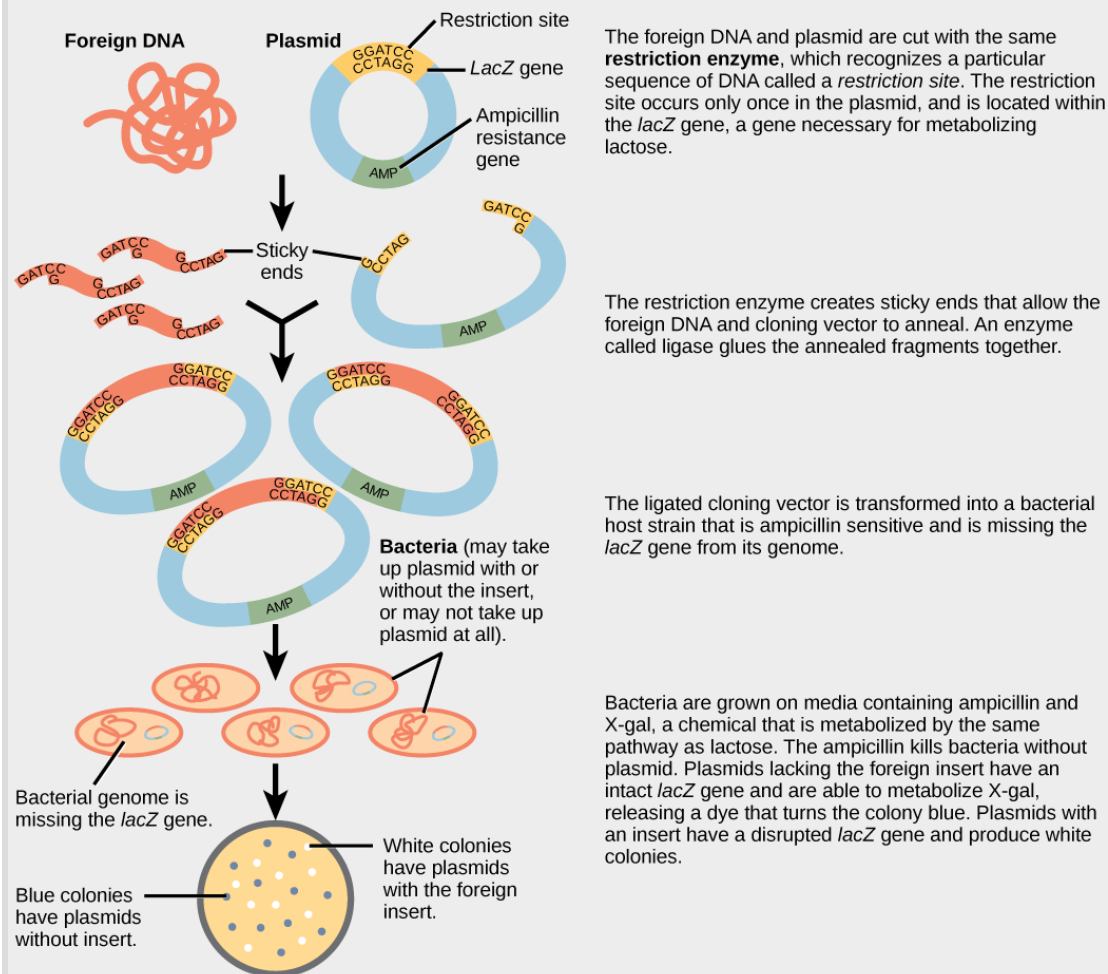
Plasmids with foreign DNA inserted into them are called **recombinant DNA** molecules because they are created artificially and do not occur in nature. They are also called chimeric molecules because the origin of different parts of the molecules can be traced back to different species of biological organisms or even to chemical synthesis. Proteins that are expressed from recombinant DNA molecules are called **recombinant**

**proteins.** Not all recombinant plasmids are capable of expressing genes. The recombinant DNA may need to be moved into a different vector (or host) that is better designed for gene expression. Plasmids may also be engineered to express proteins only when stimulated by certain environmental factors, so that scientists can control the expression of the recombinant proteins.

## Note:

### Art Connection

#### Molecular Cloning



This diagram shows the steps involved in molecular cloning.

You are working in a molecular biology lab and, unbeknownst to you, your lab partner left the foreign genomic DNA that you are planning to clone on the lab bench overnight instead of storing it in the freezer. As a result, it was degraded by nucleases, but still used in the experiment. The plasmid, on the other hand, is fine. What results would you expect from your molecular cloning experiment?

- a. There will be no colonies on the bacterial plate.
- b. There will be blue colonies only.
- c. There will be blue and white colonies.
- d. There will be white colonies only.

**Note:**

Link to Learning



View an [animation of recombination in cloning](#) from the DNA Learning Center.

## Cellular Cloning

Unicellular organisms, such as bacteria and yeast, naturally produce clones of themselves when they replicate asexually by binary fission; this is known as **cellular cloning**. The nuclear DNA duplicates by the process of mitosis, which creates an exact replica of the genetic material.

## Reproductive Cloning

**Reproductive cloning** is a method used to make a clone or an identical copy of an entire multicellular organism. Most multicellular organisms undergo reproduction by sexual means, which involves genetic hybridization of two individuals (parents), making it impossible for generation of an identical copy or a clone of either parent. Recent advances in biotechnology have made it possible to artificially induce asexual reproduction of mammals in the laboratory.

Parthenogenesis, or “virgin birth,” occurs when an embryo grows and develops without the fertilization of the egg occurring; this is a form of asexual reproduction. An example of parthenogenesis occurs in species in which the female lays an egg and if the egg is fertilized, it is a diploid egg and the individual develops into a female; if the egg is not fertilized, it remains a haploid egg and develops into a male. The unfertilized egg is called a parthenogenic, or virgin, egg. Some insects and reptiles lay parthenogenic eggs that can develop into adults.

Sexual reproduction requires two cells; when the haploid egg and sperm cells fuse, a diploid zygote results. The zygote nucleus contains the genetic information to produce a new individual. However, early embryonic development requires the cytoplasmic material contained in the egg cell. This idea forms the basis for reproductive cloning. Therefore, if the haploid nucleus of an egg cell is replaced with a diploid nucleus from the cell of any individual of the same species (called a donor), it will become a zygote that is genetically identical to the donor. Somatic cell nuclear transfer is the technique of transferring a diploid nucleus into an enucleated egg. It can be used for either therapeutic cloning or reproductive cloning.

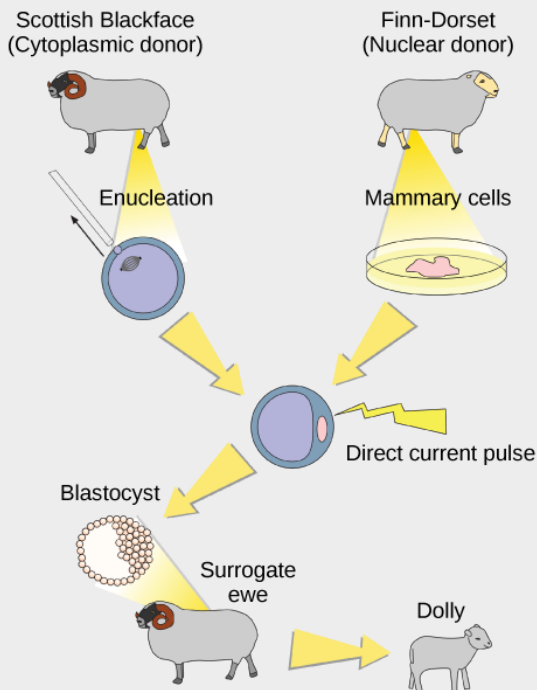
The first cloned animal was Dolly, a sheep who was born in 1996. The success rate of reproductive cloning at the time was very low. Dolly lived for seven years and died of respiratory complications ([\[link\]](#)). There is speculation that because the cell DNA belongs to an older individual, the age of the DNA may affect the life expectancy of a cloned individual. Since Dolly, several animals such as horses, bulls, and goats have been successfully cloned, although these individuals often exhibit facial, limb, and cardiac abnormalities. There have been attempts at producing cloned human embryos as sources of embryonic stem cells, sometimes referred to



as cloning for therapeutic purposes. Therapeutic cloning produces stem cells to attempt to remedy detrimental diseases or defects (unlike reproductive cloning, which aims to reproduce an organism). Still, therapeutic cloning efforts have met with resistance because of bioethical considerations.

**Note:**

**Art Connection**



Dolly the sheep was the first mammal to be cloned. To create Dolly, the nucleus was removed from a donor egg cell. The nucleus from a second sheep was then introduced into the cell, which was allowed to divide to the blastocyst stage before being implanted in a surrogate mother. (credit:

modification of work by  
"Squidonius"/Wikimedia  
Commons)

Do you think Dolly was a Finn-Dorset or a Scottish Blackface sheep?

## Genetic Engineering

**Genetic engineering** is the alteration of an organism's genotype using recombinant DNA technology to modify an organism's DNA to achieve desirable traits. The addition of foreign DNA in the form of recombinant DNA vectors generated by molecular cloning is the most common method of genetic engineering. The organism that receives the recombinant DNA is called a **genetically modified organism** (GMO). If the foreign DNA that is introduced comes from a different species, the host organism is called **transgenic**. Bacteria, plants, and animals have been genetically modified since the early 1970s for academic, medical, agricultural, and industrial purposes. In the US, GMOs such as Roundup-ready soybeans and borer-resistant corn are part of many common processed foods.

## Gene Targeting

Although classical methods of studying the function of genes began with a given phenotype and determined the genetic basis of that phenotype, modern techniques allow researchers to start at the DNA sequence level and ask: "What does this gene or DNA element do?" This technique, called reverse genetics, has resulted in reversing the classic genetic methodology. This method would be similar to damaging a body part to determine its function. An insect that loses a wing cannot fly, which means that the function of the wing is flight. The classical genetic method would compare insects that cannot fly with insects that can fly, and observe that the non-flying insects have lost wings. Similarly, mutating or deleting genes provides researchers with clues about gene function. The methods used to

disable gene function are collectively called gene targeting. **Gene targeting** is the use of recombinant DNA vectors to alter the expression of a particular gene, either by introducing mutations in a gene, or by eliminating the expression of a certain gene by deleting a part or all of the gene sequence from the genome of an organism.

## **Biotechnology in Medicine and Agriculture**

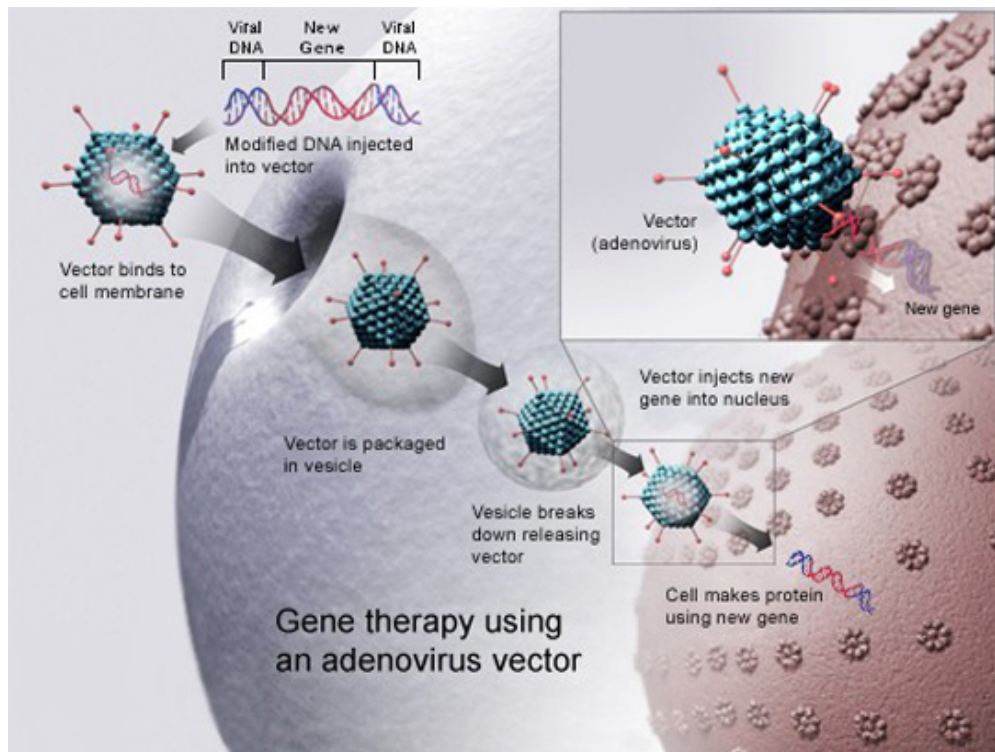
It is easy to see how biotechnology can be used for medicinal purposes. Knowledge of the genetic makeup of our species, the genetic basis of heritable diseases, and the invention of technology to manipulate and fix mutant genes provides methods to treat the disease. Biotechnology in agriculture can enhance resistance to disease, pest, and environmental stress, and improve both crop yield and quality.

## **Genetic Diagnosis and Gene Therapy**

The process of testing for suspected genetic defects before administering treatment is called **genetic diagnosis** by **genetic testing**. Depending on the inheritance patterns of a disease-causing gene, family members are advised to undergo genetic testing. For example, women diagnosed with breast cancer are usually advised to have a biopsy so that the medical team can determine the genetic basis of cancer development. Treatment plans are based on the findings of genetic tests that determine the type of cancer. If the cancer is caused by inherited gene mutations, other female relatives are also advised to undergo genetic testing and periodic screening for breast cancer. Genetic testing is also offered for fetuses (or embryos with in vitro fertilization) to determine the presence or absence of disease-causing genes in families with specific debilitating diseases.

**Gene therapy** is a genetic engineering technique used to cure disease. In its simplest form, it involves the introduction of a good gene at a random location in the genome to aid the cure of a disease that is caused by a mutated gene. The good gene is usually introduced into diseased cells as part of a vector transmitted by a virus that can infect the host cell and

deliver the foreign DNA ([link](#)). More advanced forms of gene therapy try to correct the mutation at the original site in the genome, such as is the case with treatment of severe combined immunodeficiency (SCID).



Gene therapy using an adenovirus vector can be used to cure certain genetic diseases in which a person has a defective gene. (credit: NIH)

## Production of Vaccines, Antibiotics, and Hormones

Traditional vaccination strategies use weakened or inactive forms of microorganisms to mount the initial immune response. Modern techniques use the genes of microorganisms cloned into vectors to mass produce the desired antigen. The antigen is then introduced into the body to stimulate the primary immune response and trigger immune memory. Genes cloned

from the influenza virus have been used to combat the constantly changing strains of this virus.

Antibiotics are a biotechnological product. They are naturally produced by microorganisms, such as fungi, to attain an advantage over bacterial populations. Antibiotics are produced on a large scale by cultivating and manipulating fungal cells.

Recombinant DNA technology was used to produce large-scale quantities of human insulin in *E. coli* as early as 1978. Previously, it was only possible to treat diabetes with pig insulin, which caused allergic reactions in humans because of differences in the gene product. In addition, human growth hormone (HGH) is used to treat growth disorders in children. The HGH gene was cloned from a cDNA library and inserted into *E. coli* cells by cloning it into a bacterial vector.

## **Transgenic Animals**

Although several recombinant proteins used in medicine are successfully produced in bacteria, some proteins require a eukaryotic animal host for proper processing. For this reason, the desired genes are cloned and expressed in animals, such as sheep, goats, chickens, and mice. Animals that have been modified to express recombinant DNA are called transgenic animals. Several human proteins are expressed in the milk of transgenic sheep and goats, and some are expressed in the eggs of chickens. Mice have been used extensively for expressing and studying the effects of recombinant genes and mutations.

## **Transgenic Plants**

Manipulating the DNA of plants (i.e., creating GMOs) has helped to create desirable traits, such as disease resistance, herbicide and pesticide resistance, better nutritional value, and better shelf-life ([\[link\]](#)). Plants are the most important source of food for the human population. Farmers developed ways to select for plant varieties with desirable traits long before modern-day biotechnology practices were established.



Corn, a major agricultural crop used to create products for a variety of industries, is often modified through plant biotechnology. (credit: Keith Weller, USDA)

Plants that have received recombinant DNA from other species are called transgenic plants. Because they are not natural, transgenic plants and other GMOs are closely monitored by government agencies to ensure that they are fit for human consumption and do not endanger other plant and animal life. Because foreign genes can spread to other species in the environment, extensive testing is required to ensure ecological stability. Staples like corn, potatoes, and tomatoes were the first crop plants to be genetically engineered.

## **Transformation of Plants Using *Agrobacterium tumefaciens***

Gene transfer occurs naturally between species in microbial populations. Many viruses that cause human diseases, such as cancer, act by incorporating their DNA into the human genome. In plants, tumors caused by the bacterium *Agrobacterium tumefaciens* occur by transfer of DNA from the bacterium to the plant. Although the tumors do not kill the plants, they make the plants stunted and more susceptible to harsh environmental conditions. Many plants, such as walnuts, grapes, nut trees, and beets, are affected by *A. tumefaciens*. The artificial introduction of DNA into plant cells is more challenging than in animal cells because of the thick plant cell wall.

Researchers used the natural transfer of DNA from *Agrobacterium* to a plant host to introduce DNA fragments of their choice into plant hosts. In nature, the disease-causing *A. tumefaciens* have a set of plasmids, called the **Ti plasmids** (tumor-inducing plasmids), that contain genes for the production of tumors in plants. DNA from the Ti plasmid integrates into the infected plant cell's genome. Researchers manipulate the Ti plasmids to remove the tumor-causing genes and insert the desired DNA fragment for transfer into the plant genome. The Ti plasmids carry antibiotic resistance genes to aid selection and can be propagated in *E. coli* cells as well.

## **The Organic Insecticide *Bacillus thuringiensis***

*Bacillus thuringiensis* (Bt) is a bacterium that produces protein crystals during sporulation that are toxic to many insect species that affect plants. Bt toxin has to be ingested by insects for the toxin to be activated. Insects that have eaten Bt toxin stop feeding on the plants within a few hours. After the toxin is activated in the intestines of the insects, death occurs within a couple of days. Modern biotechnology has allowed plants to encode their own crystal Bt toxin that acts against insects. The crystal toxin genes have been cloned from Bt and introduced into plants. Bt toxin has been found to be safe for the environment, non-toxic to humans and other mammals, and is approved for use by organic farmers as a natural insecticide.

## **Flavr Savr Tomato**

The first GM crop to be introduced into the market was the Flavr Savr Tomato produced in 1994. Antisense RNA technology was used to slow down the process of softening and rotting caused by fungal infections, which led to increased shelf life of the GM tomatoes. Additional genetic modification improved the flavor of this tomato. The Flavr Savr tomato did not successfully stay in the market because of problems maintaining and shipping the crop.

## **Section Summary**

Nucleic acids can be isolated from cells for the purposes of further analysis by breaking open the cells and enzymatically destroying all other major macromolecules. Fragmented or whole chromosomes can be separated on the basis of size by gel electrophoresis. Short stretches of DNA or RNA can be amplified by PCR. Southern and northern blotting can be used to detect the presence of specific short sequences in a DNA or RNA sample. The term “cloning” may refer to cloning small DNA fragments (molecular cloning), cloning cell populations (cellular cloning), or cloning entire organisms (reproductive cloning). Genetic testing is performed to identify disease-causing genes, and gene therapy is used to cure an inheritable disease.

Transgenic organisms possess DNA from a different species, usually generated by molecular cloning techniques. Vaccines, antibiotics, and hormones are examples of products obtained by recombinant DNA technology. Transgenic plants are usually created to improve characteristics of crop plants.

## **Art Connections**

### **Exercise:**



**Problem:**

[\[link\]](#) You are working in a molecular biology lab and, unbeknownst to you, your lab partner left the foreign genomic DNA that you are planning to clone on the lab bench overnight instead of storing it in the freezer. As a result, it was degraded by nucleases, but still used in the experiment. The plasmid, on the other hand, is fine. What results would you expect from your molecular cloning experiment?

- a. There will be no colonies on the bacterial plate.
- b. There will be blue colonies only.
- c. There will be blue and white colonies.
- d. There will be white colonies only.

---

**Solution:**

[\[link\]](#) B. The experiment would result in blue colonies only.

**Exercise:****Problem:**

[\[link\]](#) Do you think Dolly was a Finn-Dorset or a Scottish Blackface sheep?

---

**Solution:**

[\[link\]](#) Dolly was a Finn-Dorset sheep because even though the original cell came from a Scottish blackface sheep and the surrogate mother was a Scottish blackface, the DNA came from a Finn-Dorset.

**Review Questions****Exercise:**

**Problem:** GMOs are created by \_\_\_\_\_.

- a. generating genomic DNA fragments with restriction endonucleases
- b. introducing recombinant DNA into an organism by any means
- c. overexpressing proteins in *E. coli*.
- d. all of the above

---

**Solution:**

B

**Exercise:**

**Problem:**

Gene therapy can be used to introduce foreign DNA into cells \_\_\_\_\_.

- a. for molecular cloning
- b. by PCR
- c. of tissues to cure inheritable disease
- d. all of the above

---

**Solution:**

C

**Exercise:**

**Problem:** Insulin produced by molecular cloning:

- a. is of pig origin
- b. is a recombinant protein
- c. is made by the human pancreas
- d. is recombinant DNA

---

**Solution:**

B

**Exercise:**

**Problem:**Bt toxin is considered to be \_\_\_\_\_.

- a. a gene for modifying insect DNA
  - b. an organic insecticide produced by bacteria
  - c. useful for humans to fight against insects
  - d. a recombinant protein
- 

**Solution:**

B

**Exercise:**

**Problem:**The Flavr Savr Tomato:

- a. is a variety of vine-ripened tomato in the supermarket
  - b. was created to have better flavor and shelf-life
  - c. does not undergo soft rot
  - d. all of the above
- 

**Solution:**

D

**Free Response**

**Exercise:**

**Problem:**Describe the process of Southern blotting.

---

**Solution:**

Southern blotting is the transfer of DNA that has been enzymatically cut into fragments and run on an agarose gel onto a nylon membrane. The DNA fragments that are on the nylon membrane can be denatured to make them single-stranded, and then probed with small DNA fragments that are radioactively or fluorescently labeled, to detect the presence of specific sequences. An example of the use of Southern blotting would be in analyzing the presence, absence, or variation of a disease gene in genomic DNA from a group of patients.

**Exercise:**

**Problem:**

A researcher wants to study cancer cells from a patient with breast cancer. Is cloning the cancer cells an option?

---

**Solution:**

Cellular cloning of the breast cancer cells will establish a cell line, which can be used for further analysis

**Exercise:**

**Problem:**

How would a scientist introduce a gene for herbicide resistance into a plant?

---

**Solution:**

By identifying an herbicide resistance gene and cloning it into a plant expression vector system, like the Ti plasmid system from *Agrobacterium tumefaciens*. The scientist would then introduce it into the plant cells by transformation, and select cells that have taken up and integrated the herbicide-resistance gene into the genome.

**Exercise:**

**Problem:**

If you had a chance to get your genome sequenced, what are some questions you might be able to have answered about yourself?

---

**Solution:**

What diseases am I prone to and what precautions should I take? Am I a carrier for any disease-causing genes that may be passed on to children?

**Glossary**

antibiotic resistance

ability of an organism to be unaffected by the actions of an antibiotic

biotechnology

use of biological agents for technological advancement

cellular cloning

production of identical cell populations by binary fission

clone

exact replica

foreign DNA

DNA that belongs to a different species or DNA that is artificially synthesized

gel electrophoresis

technique used to separate molecules on the basis of size using electric charge

gene targeting

method for altering the sequence of a specific gene by introducing the modified version on a vector

gene therapy

technique used to cure inheritable diseases by replacing mutant genes with good genes

genetic diagnosis

diagnosis of the potential for disease development by analyzing disease-causing genes

genetic engineering

alteration of the genetic makeup of an organism

genetic testing

process of testing for the presence of disease-causing genes

genetically modified organism (GMO)

organism whose genome has been artificially changed

host DNA

DNA that is present in the genome of the organism of interest

lysis buffer

solution used to break the cell membrane and release cell contents

molecular cloning

cloning of DNA fragments

multiple cloning site (MCS)

site that can be recognized by multiple restriction endonucleases

northern blotting

transfer of RNA from a gel to a nylon membrane

polymerase chain reaction (PCR)

technique used to amplify DNA

probe

small DNA fragment used to determine if the complementary sequence is present in a DNA sample

protease

enzyme that breaks down proteins

recombinant DNA

combination of DNA fragments generated by molecular cloning that does not exist in nature; also known as a chimeric molecule

recombinant protein

protein product of a gene derived by molecular cloning

reproductive cloning

cloning of entire organisms

restriction endonuclease

enzyme that can recognize and cleave specific DNA sequences

reverse genetics

method of determining the function of a gene by starting with the gene itself instead of starting with the gene product

reverse transcriptase PCR (RT-PCR)

PCR technique that involves converting RNA to DNA by reverse transcriptase

ribonuclease

enzyme that breaks down RNA

Southern blotting

transfer of DNA from a gel to a nylon membrane

Ti plasmid

plasmid system derived from *Agrobacterium tumifaciens* that has been used by scientists to introduce foreign DNA into plant cells

transgenic

organism that receives DNA from a different species

## Mapping Genomes

By the end of this section, you will be able to:

- Define genomics
- Describe genetic and physical maps
- Describe genomic mapping methods

**Genomics** is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. **Genome mapping** is the process of finding the locations of genes on each chromosome. The maps created by genome mapping are comparable to the maps that we use to navigate streets. A **genetic map** is an illustration that lists genes and their location on a chromosome. Genetic maps provide the big picture (similar to a map of interstate highways) and use genetic markers (similar to landmarks). A **genetic marker** is a gene or sequence on a chromosome that co-segregates (shows genetic linkage) with a specific trait. Early geneticists called this linkage analysis. Physical maps present the intimate details of smaller regions of the chromosomes (similar to a detailed road map). A **physical map** is a representation of the physical distance, in nucleotides, between genes or genetic markers. Both genetic linkage maps and physical maps are required to build a complete picture of the genome. Having a complete map of the genome makes it easier for researchers to study individual genes. Human genome maps help researchers in their efforts to identify human disease-causing genes related to illnesses like cancer, heart disease, and cystic fibrosis. Genome mapping can be used in a variety of other applications, such as using live microbes to clean up pollutants or even prevent pollution. Research involving plant genome mapping may lead to producing higher crop yields or developing plants that better adapt to climate change.

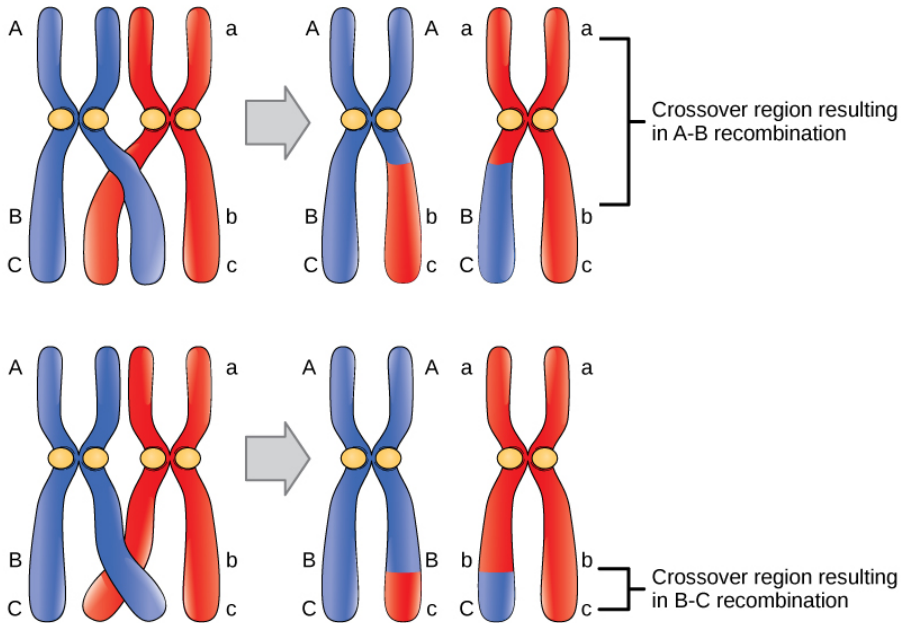
## Genetic Maps

The study of genetic maps begins with **linkage analysis**, a procedure that analyzes the recombination frequency between genes to determine if they are linked or show independent assortment. The term *linkage* was used before the discovery of DNA. Early geneticists relied on the observation of



phenotypic changes to understand the genotype of an organism. Shortly after Gregor Mendel (the father of modern genetics) proposed that traits were determined by what are now known as genes, other researchers observed that different traits were often inherited together, and thereby deduced that the genes were physically linked by being located on the same chromosome. The mapping of genes relative to each other based on linkage analysis led to the development of the first genetic maps.

Observations that certain traits were always linked and certain others were not linked came from studying the offspring of crosses between parents with different traits. For example, in experiments performed on the garden pea, it was discovered that the color of the flower and shape of the plant's pollen were linked traits, and therefore the genes encoding these traits were in close proximity on the same chromosome. The exchange of DNA between homologous pairs of chromosomes is called **genetic recombination**, which occurs by the crossing over of DNA between homologous strands of DNA, such as nonsister chromatids. Linkage analysis involves studying the recombination frequency between any two genes. The greater the distance between two genes, the higher the chance that a recombination event will occur between them, and the higher the recombination frequency between them. Two possibilities for recombination between two nonsister chromatids during meiosis are shown in [\[link\]](#). If the recombination frequency between two genes is less than 50 percent, they are said to be linked.



Crossover may occur at different locations on the chromosome. Recombination between genes *A* and *B* is more frequent than recombination between genes *B* and *C* because genes *A* and *B* are farther apart; a crossover is therefore more likely to occur between them.

The generation of genetic maps requires markers, just as a road map requires landmarks (such as rivers and mountains). Early genetic maps were based on the use of known genes as markers. More sophisticated markers, including those based on non-coding DNA, are now used to compare the genomes of individuals in a population. Although individuals of a given species are genetically similar, they are not identical; every individual has a unique set of traits. These minor differences in the genome between individuals in a population are useful for the purposes of genetic mapping. In general, a good genetic marker is a region on the chromosome that shows variability or polymorphism (multiple forms) in the population.

Some genetic markers used in generating genetic maps are **restriction fragment length polymorphisms (RFLP)**, variable number of tandem

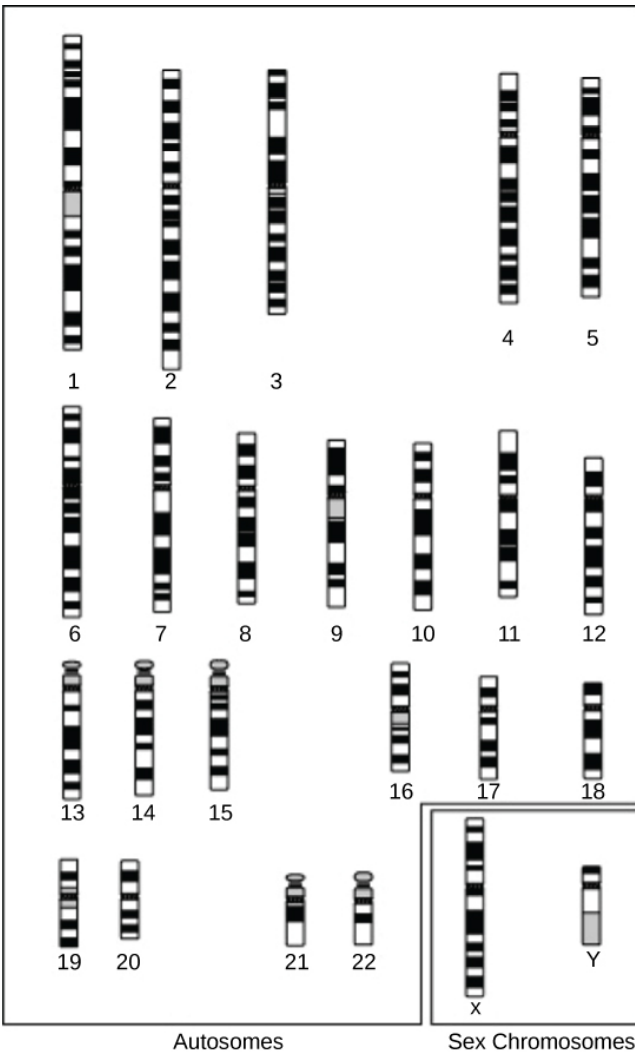
repeats (VNTRs), **microsatellite polymorphisms**, and the **single nucleotide polymorphisms** (SNPs). RFLPs (sometimes pronounced “rif-lips”) are detected when the DNA of an individual is cut with a restriction endonuclease that recognizes specific sequences in the DNA to generate a series of DNA fragments, which are then analyzed by gel electrophoresis. The DNA of every individual will give rise to a unique pattern of bands when cut with a particular set of restriction endonucleases; this is sometimes referred to as an individual’s DNA “fingerprint.” Certain regions of the chromosome that are subject to polymorphism will lead to the generation of the unique banding pattern. VNTRs are repeated sets of nucleotides present in the non-coding regions of DNA. Non-coding, or “junk,” DNA has no known biological function; however, research shows that much of this DNA is actually transcribed. While its function is uncertain, it is certainly active, and it may be involved in the regulation of coding genes. The number of repeats may vary in individual organisms of a population. Microsatellite polymorphisms are similar to VNTRs, but the repeat unit is very small. SNPs are variations in a single nucleotide.

Because genetic maps rely completely on the natural process of recombination, mapping is affected by natural increases or decreases in the level of recombination in any given area of the genome. Some parts of the genome are recombination hotspots, whereas others do not show a propensity for recombination. For this reason, it is important to look at mapping information developed by multiple methods.

## Physical Maps

A physical map provides detail of the actual physical distance between genetic markers, as well as the number of nucleotides. There are three methods used to create a physical map: cytogenetic mapping, radiation hybrid mapping, and sequence mapping. **Cytogenetic mapping** uses information obtained by microscopic analysis of stained sections of the chromosome ([\[link\]](#)). It is possible to determine the approximate distance between genetic markers using cytogenetic mapping, but not the exact distance (number of base pairs). **Radiation hybrid mapping** uses radiation, such as x-rays, to break the DNA into fragments. The amount of radiation can be adjusted to create smaller or larger fragments. This technique

overcomes the limitation of genetic mapping and is not affected by increased or decreased recombination frequency. **Sequence mapping** resulted from DNA sequencing technology that allowed for the creation of detailed physical maps with distances measured in terms of the number of base pairs. The creation of **genomic libraries** and **complementary DNA (cDNA) libraries** (collections of cloned sequences or all DNA from a genome) has sped up the process of physical mapping. A genetic site used to generate a physical map with sequencing technology (a sequence-tagged site, or STS) is a unique sequence in the genome with a known exact chromosomal location. An **expressed sequence tag (EST)** and a single sequence length polymorphism (SSLP) are common STSs. An EST is a short STS that is identified with cDNA libraries, while SSLPs are obtained from known genetic markers and provide a link between genetic maps and physical maps.



A cytogenetic map shows the appearance of a chromosome after it is stained and examined under a microscope. (credit: National Human Genome Research Institute)

## Integration of Genetic and Physical Maps

Genetic maps provide the outline and physical maps provide the details. It is easy to understand why both types of genome mapping techniques are

important to show the big picture. Information obtained from each technique is used in combination to study the genome. Genomic mapping is being used with different model organisms that are used for research. Genome mapping is still an ongoing process, and as more advanced techniques are developed, more advances are expected. Genome mapping is similar to completing a complicated puzzle using every piece of available data. Mapping information generated in laboratories all over the world is entered into central databases, such as GenBank at the National Center for Biotechnology Information (NCBI). Efforts are being made to make the information more easily accessible to researchers and the general public. Just as we use global positioning systems instead of paper maps to navigate through roadways, NCBI has created a genome viewer tool to simplify the data-mining process.

**Note:**

Scientific Method Connection

**How to Use a Genome Map Viewer**

Problem statement: Do the human, macaque, and mouse genomes contain common DNA sequences?

Develop a hypothesis.

To test the hypothesis, click this [link](#).

In Search box on the left panel, type any gene name or phenotypic characteristic, such as iris pigmentation (eye color). Select the species you want to study, and then press Enter. The genome map viewer will indicate which chromosome encodes the gene in your search. Click each hit in the genome viewer for more detailed information. This type of search is the most basic use of the genome viewer; it can also be used to compare sequences between species, as well as many other complicated tasks.

Is the hypothesis correct? Why or why not?

**Note:**

Link to Learning



Online Mendelian Inheritance in Man (OMIM) is a searchable online catalog of human genes and genetic disorders. This website shows genome mapping information, and also details the history and research of each trait and disorder. Click this [link](#) to search for traits (such as handedness) and genetic disorders (such as diabetes).

## Section Summary

Genome mapping is similar to solving a big, complicated puzzle with pieces of information coming from laboratories all over the world. Genetic maps provide an outline for the location of genes within a genome, and they estimate the distance between genes and genetic markers on the basis of recombination frequencies during meiosis. Physical maps provide detailed information about the physical distance between the genes. The most detailed information is available through sequence mapping. Information from all mapping and sequencing sources is combined to study an entire genome.

## Review Questions

### Exercise:

**Problem:** ESTs are \_\_\_\_\_.

- a. generated after a cDNA library is made
- b. unique sequences in the genome
- c. useful for mapping using sequence information
- d. all of the above

---

**Solution:**

D

**Exercise:**

**Problem:** Linkage analysis \_\_\_\_\_.

- a. is used to create a physical map
- b. is based on the natural recombination process
- c. requires radiation hybrid mapping
- d. involves breaking and re-joining of DNA artificially

---

**Solution:**

B

**Exercise:**

**Problem:** Genetic recombination occurs by which process?

- a. independent assortment
- b. crossing over
- c. chromosome segregation
- d. sister chromatids

---

**Solution:**

B

**Exercise:**

**Problem:** Individual genetic maps in a given species are:

- a. genetically similar
- b. genetically identical



- c. genetically dissimilar
  - d. not useful in species analysis
- 

**Solution:**

A

**Exercise:**

**Problem:**

Information obtained by microscopic analysis of stained chromosomes is used in:

- a. radiation hybrid mapping
  - b. sequence mapping
  - c. RFLP mapping
  - d. cytogenetic mapping
- 

**Solution:**

D

**Free Response**

**Exercise:**

**Problem:**

Why is so much effort being poured into genome mapping applications?

---

**Solution:**

Genome mapping has many different applications and provides comprehensive information that can be used for predictive purposes.

**Exercise:**

**Problem:**

How could a genetic map of the human genome help find a cure for cancer?

---

**Solution:**

A human genetic map can help identify genetic markers and sequences associated with high cancer risk, which can help to screen and provide early detection of different types of cancer.

**Glossary**

cytogenetic mapping

technique that uses a microscope to create a map from stained chromosomes

expressed sequence tag (EST)

short STS that is identified with cDNA

genetic map

outline of genes and their location on a chromosome

genetic marker

gene or sequence on a chromosome with a known location that is associated with a specific trait

genetic recombination

exchange of DNA between homologous pairs of chromosomes

genome mapping

process of finding the location of genes on each chromosome

cDNA library

collection of cloned cDNA sequences

genomic library

collection of cloned DNA which represents all of the sequences and fragments from a genome

genomics

study of entire genomes including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species

linkage analysis

procedure that analyzes the recombination of genes to determine if they are linked

microsatellite polymorphism

variation between individuals in the sequence and number of repeats of microsatellite DNA

physical map

representation of the physical distance between genes or genetic markers

radiation hybrid mapping

information obtained by fragmenting the chromosome with x-rays

restriction fragment length polymorphism (RFLP)

variation between individuals in the length of DNA fragments generated by restriction endonucleases

sequence mapping

mapping information obtained after DNA sequencing

single nucleotide polymorphism (SNP)

variation between individuals in a single nucleotide

variable number of tandem repeats (VNTRs)

variation in the number of tandem repeats between individuals in the population

## Whole-Genome Sequencing

By the end of this section, you will be able to:

- Describe three types of sequencing
- Define whole-genome sequencing

Although there have been significant advances in the medical sciences in recent years, doctors are still confounded by some diseases, and they are using whole-genome sequencing to get to the bottom of the problem.

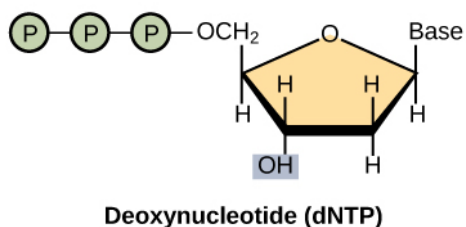
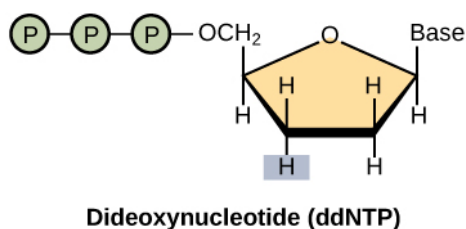
**Whole-genome sequencing** is a process that determines the DNA sequence of an entire genome. Whole-genome sequencing is a brute-force approach to problem solving when there is a genetic basis at the core of a disease. Several laboratories now provide services to sequence, analyze, and interpret entire genomes.

For example, whole-exome sequencing is a lower-cost alternative to whole genome sequencing. In exome sequencing, only the coding, exon-producing regions of the DNA are sequenced. In 2010, whole-exome sequencing was used to save a young boy whose intestines had multiple mysterious abscesses. The child had several colon operations with no relief. Finally, whole-exome sequencing was performed, which revealed a defect in a pathway that controls apoptosis (programmed cell death). A bone-marrow transplant was used to overcome this genetic disorder, leading to a cure for the boy. He was the first person to be successfully treated based on a diagnosis made by whole-exome sequencing. Today, human genome sequencing is more readily available and can be completed in a day or two for about \$1000.

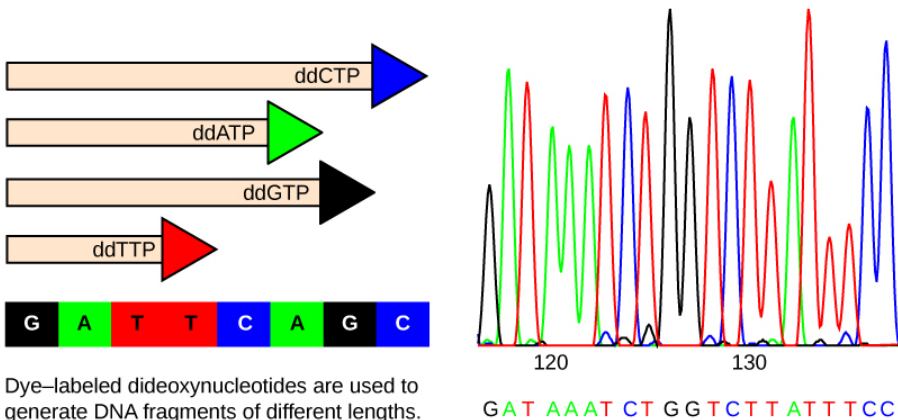
## Strategies Used in Sequencing Projects

The basic sequencing technique used in all modern day sequencing projects is the chain termination method (also known as the dideoxy method), which was developed by Fred Sanger in the 1970s. The chain termination method involves DNA replication of a single-stranded template with the use of a primer and a regular **deoxynucleotide** (dNTP), which is a monomer, or a single unit, of DNA. The primer and dNTP are mixed with a small proportion of fluorescently labeled **dideoxynucleotides** (ddNTPs). The

ddNTPs are monomers that are missing a hydroxyl group ( $\text{-OH}$ ) at the site at which another nucleotide usually attaches to form a chain ([\[link\]](#)). Each ddNTP is labeled with a different color of fluorophore. Every time a ddNTP is incorporated in the growing complementary strand, it terminates the process of DNA replication, which results in multiple short strands of replicated DNA that are each terminated at a different point during replication. When the reaction mixture is processed by gel electrophoresis after being separated into single strands, the multiple newly replicated DNA strands form a ladder because of the differing sizes. Because the ddNTPs are fluorescently labeled, each band on the gel reflects the size of the DNA strand and the ddNTP that terminated the reaction. The different colors of the fluorophore-labeled ddNTPs help identify the ddNTP incorporated at that position. Reading the gel on the basis of the color of each band on the ladder produces the sequence of the template strand ([\[link\]](#)).



A dideoxynucleotide is similar in structure to a deoxynucleotide, but is missing the 3' hydroxyl group (indicated by the box). When a dideoxynucleotide is incorporated into a DNA strand, DNA synthesis stops.



Frederick Sanger's dideoxy chain termination method is illustrated. Using dideoxynucleotides, the DNA fragment can be terminated at different points. The DNA is separated on the basis of size, and these bands, based on the size of the fragments, can be read.

## Early Strategies: Shotgun Sequencing and Pair-Wise End Sequencing

In **shotgun sequencing** method, several copies of a DNA fragment are cut randomly into many smaller pieces (somewhat like what happens to a round shot cartridge when fired from a shotgun). All of the segments are then sequenced using the chain-sequencing method. Then, with the help of a computer, the fragments are analyzed to see where their sequences overlap. By matching up overlapping sequences at the end of each fragment, the entire DNA sequence can be reformed. A larger sequence that is assembled from overlapping shorter sequences is called a **contig**. As an analogy, consider that someone has four copies of a landscape photograph that you have never seen before and know nothing about how it should appear. The person then rips up each photograph with their hands, so that different size pieces are present from each copy. The person then mixes all of the pieces together and asks you to reconstruct the photograph. In one of the smaller pieces you see a mountain. In a larger piece, you see that the same mountain

is behind a lake. A third fragment shows only the lake, but it reveals that there is a cabin on the shore of the lake. Therefore, from looking at the overlapping information in these three fragments, you know that the picture contains a mountain behind a lake that has a cabin on its shore. This is the principle behind reconstructing entire DNA sequences using shotgun sequencing.

Originally, shotgun sequencing only analyzed one end of each fragment for overlaps. This was sufficient for sequencing small genomes. However, the desire to sequence larger genomes, such as that of a human, led to the development of double-barrel shotgun sequencing, more formally known as **pairwise-end sequencing**. In pairwise-end sequencing, both ends of each fragment are analyzed for overlap. Pairwise-end sequencing is, therefore, more cumbersome than shotgun sequencing, but it is easier to reconstruct the sequence because there is more available information.

## Next-generation Sequencing

Since 2005, automated sequencing techniques used by laboratories are under the umbrella of **next-generation sequencing**, which is a group of automated techniques used for rapid DNA sequencing. These automated low-cost sequencers can generate sequences of hundreds of thousands or millions of short fragments (25 to 500 base pairs) in the span of one day. These sequencers use sophisticated software to get through the cumbersome process of putting all the fragments in order.

### Note:

#### Evolution Connection

#### Comparing Sequences

A sequence alignment is an arrangement of proteins, DNA, or RNA; it is used to identify regions of similarity between cell types or species, which may indicate conservation of function or structures. Sequence alignments may be used to construct phylogenetic trees. The following website uses a software program called [BLAST \(basic local alignment search tool\)](#).

Under “Basic Blast,” click “Nucleotide Blast.” Input the following sequence into the large "query sequence" box: ATTGCTTCGATTGCA. Below the box, locate the "Species" field and type "human" or "Homo sapiens". Then click “BLAST” to compare the inputted sequence against known sequences of the human genome. The result is that this sequence occurs in over a hundred places in the human genome. Scroll down below the graphic with the horizontal bars and you will see short description of each of the matching hits. Pick one of the hits near the top of the list and click on "Graphics". This will bring you to a page that shows where the sequence is found within the entire human genome. You can move the slider that looks like a green flag back and forth to view the sequences immediately around the selected gene. You can then return to your selected sequence by clicking the "ATG" button.

## Use of Whole-Genome Sequences of Model Organisms

The first genome to be completely sequenced was of a bacterial virus, the bacteriophage *φx174* (5368 base pairs); this was accomplished by Fred Sanger using shotgun sequencing. Several other organelle and viral genomes were later sequenced. The first organism whose genome was sequenced was the bacterium *Haemophilus influenzae*; this was accomplished by Craig Venter in the 1980s. Approximately 74 different laboratories collaborated on the sequencing of the genome of the yeast *Saccharomyces cerevisiae*, which began in 1989 and was completed in 1996, because it was 60 times bigger than any other genome that had been sequenced. By 1997, the genome sequences of two important model organisms were available: the bacterium *Escherichia coli* K12 and the yeast *Saccharomyces cerevisiae*. Genomes of other model organisms, such as the mouse *Mus musculus*, the fruit fly *Drosophila melanogaster*, the nematode *Caenorhabditis. elegans*, and humans *Homo sapiens* are now known. A lot of basic research is performed in model organisms because the information can be applied to genetically similar organisms. A **model organism** is a species that is studied as a model to understand the biological processes in other species represented by the model organism. Having entire genomes sequenced helps with the research efforts in these model organisms. The



process of attaching biological information to gene sequences is called **genome annotation**. Annotation of gene sequences helps with basic experiments in molecular biology, such as designing PCR primers and RNA targets.

**Note:**

Link to Learning



Click through each step of genome sequencing at this [site](#).

## Uses of Genome Sequences

**DNA microarrays** are methods used to detect gene expression by analyzing an array of DNA fragments that are fixed to a glass slide or a silicon chip to identify active genes and identify sequences. Almost one million genotypic abnormalities can be discovered using microarrays, whereas whole-genome sequencing can provide information about all six billion base pairs in the human genome. Although the study of medical applications of genome sequencing is interesting, this discipline tends to dwell on abnormal gene function. Knowledge of the entire genome will allow future onset diseases and other genetic disorders to be discovered early, which will allow for more informed decisions to be made about lifestyle, medication, and having children. Genomics is still in its infancy, although someday it may become routine to use whole-genome sequencing to screen every newborn to detect genetic abnormalities.

In addition to disease and medicine, genomics can contribute to the development of novel enzymes that convert biomass to biofuel, which

results in higher crop and fuel production, and lower cost to the consumer. This knowledge should allow better methods of control over the microbes that are used in the production of biofuels. Genomics could also improve the methods used to monitor the impact of pollutants on ecosystems and help clean up environmental contaminants. Genomics has allowed for the development of agrochemicals and pharmaceuticals that could benefit medical science and agriculture.

It sounds great to have all the knowledge we can get from whole-genome sequencing; however, humans have a responsibility to use this knowledge wisely. Otherwise, it could be easy to misuse the power of such knowledge, leading to discrimination based on a person's genetics, human genetic engineering, and other ethical concerns. This information could also lead to legal issues regarding health and privacy.

## **Section Summary**

Whole-genome sequencing is the latest available resource to treat genetic diseases. Some doctors are using whole-genome sequencing to save lives. Genomics has many industrial applications including biofuel development, agriculture, pharmaceuticals, and pollution control. The basic principle of all modern-day sequencing strategies involves the chain termination method of sequencing.

Although the human genome sequences provide key insights to medical professionals, researchers use whole-genome sequences of model organisms to better understand the genome of the species. Automation and the decreased cost of whole-genome sequencing may lead to personalized medicine in the future.

## **Review Questions**

### **Exercise:**

**Problem:** The chain termination method of sequencing:

- a. uses labeled ddNTPs
  - b. uses only dideoxynucleotides
  - c. uses only deoxynucleotides
  - d. uses labeled dNTPs
- 

**Solution:**

A

**Exercise:**

**Problem:** Whole-genome sequencing can be used for advances in:

- a. the medical field
  - b. agriculture
  - c. biofuels
  - d. all of the above
- 

**Solution:**

D

**Exercise:**

**Problem:** Sequencing an individual person's genome

- a. is currently possible
  - b. could lead to legal issues regarding discrimination and privacy
  - c. could help make informed choices about medical treatment
  - d. all of the above
- 

**Solution:**

D

**Exercise:**

**Problem:**

What is the most challenging issue facing genome sequencing?

- a. the inability to develop fast and accurate sequencing techniques
- b. the ethics of using information from genomes at the individual level
- c. the availability and stability of DNA
- d. all of the above

---

**Solution:**

B

**Glossary**

chain termination method

method of DNA sequencing using labeled dideoxynucleotides to terminate DNA replication; it is also called the dideoxy method or the Sanger method

contig

larger sequence of DNA assembled from overlapping shorter sequences

deoxynucleotide

individual monomer (single unit) of DNA

dideoxynucleotide

individual monomer of DNA that is missing a hydroxyl group (–OH)

DNA microarray

method used to detect gene expression by analyzing an array of DNA fragments that are fixed to a glass slide or a silicon chip to identify active genes and identify sequences

genome annotation

process of attaching biological information to gene sequences

model organism

species that is studied and used as a model to understand the biological processes in other species represented by the model organism

next-generation sequencing

group of automated techniques used for rapid DNA sequencing

shotgun sequencing

method used to sequence multiple DNA fragments to generate the sequence of a large piece of DNA

whole-genome sequencing

process that determines the DNA sequence of an entire genome

## Applying Genomics

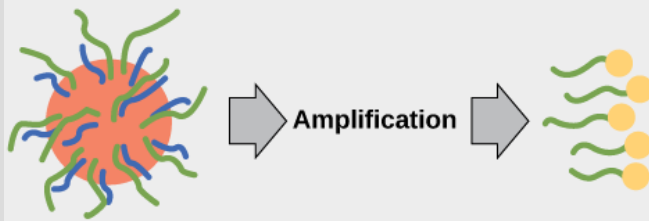
By the end of this section, you will be able to:

- Explain pharmacogenomics
- Define polygenic

The introduction of DNA sequencing and whole genome sequencing projects, particularly the Human Genome project, has expanded the applicability of DNA sequence information. Genomics is now being used in a wide variety of fields, such as metagenomics, pharmacogenomics, and mitochondrial genomics. The most commonly known application of genomics is to understand and find cures for diseases.

## Predicting Disease Risk at the Individual Level

Predicting the risk of disease involves screening currently healthy individuals by genome analysis at the individual level. Intervention with lifestyle changes and drugs can be recommended before disease onset. However, this approach is most applicable when the problem resides within a single gene defect. Such defects only account for approximately 5 percent of diseases in developed countries. Most of the common diseases, such as heart disease, are multi-factored or **polygenic**, which is a phenotypic characteristic that involves two or more genes, and also involve environmental factors such as diet. In April 2010, scientists at Stanford University published the genome analysis of a healthy individual (Stephen Quake, a scientist at Stanford University, who had his genome sequenced); the analysis predicted his propensity to acquire various diseases. A risk assessment was performed to analyze Quake's percentage of risk for 55 different medical conditions. A rare genetic mutation was found, which showed him to be at risk for sudden heart attack. He was also predicted to have a 23 percent risk of developing prostate cancer and a 1.4 percent risk of developing Alzheimer's. The scientists used databases and several publications to analyze the genomic data. Even though genomic sequencing is becoming more affordable and analytical tools are becoming more reliable, ethical issues surrounding genomic analysis at a population level remain to be addressed.

**Note:****Art Connection****PCA3**

**Step 1:**  
*PCA3* mRNA  
anneals to  
complementary  
DNA primers  
that are attached  
to beads.

**Step 2:**  
The mRNA is  
amplified using  
reverse-  
transcriptase  
PCR.

**Step 3:**  
The mRNA is  
detected using  
a chemilumi-  
nescent probe.

*PCA3* is a gene that is expressed in prostate epithelial cells and overexpressed in cancerous cells. A high concentration of *PCA3* in urine is indicative of prostate cancer. The *PCA3* test is considered to be a better indicator of cancer than the more well know PSA test, which measures the level of PSA (prostate-specific antigen) in the blood.

In 2011, the United States Preventative Services Task Force recommended against using the PSA test to screen healthy men for prostate cancer. Their recommendation is based on evidence that screening does not reduce the risk of death from prostate cancer. Prostate cancer often develops very slowly and does not cause problems, while the cancer treatment can have severe side effects. The *PCA3* test is considered to be more accurate, but screening may still result in men who would not have been harmed by the cancer itself suffering side effects from treatment. What do you think? Should all healthy men be screened for prostate cancer using the *PCA3* or

PSA test? Should people in general be screened to find out if they have a genetic risk for cancer or other diseases?

## Pharmacogenomics and Toxicogenomics

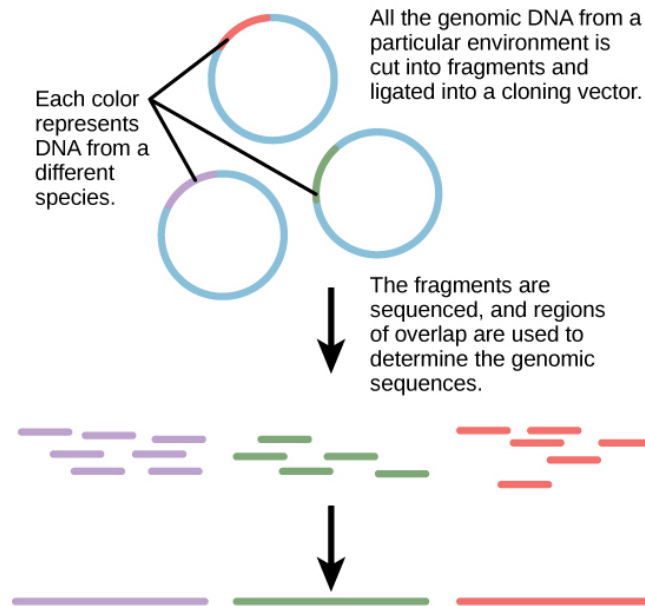
**Pharmacogenomics**, also called toxicogenomics, involves evaluating the effectiveness and safety of drugs on the basis of information from an individual's genomic sequence. Genomic responses to drugs can be studied using experimental animals (such as laboratory rats or mice) or live cells in the laboratory before embarking on studies with humans. Studying changes in gene expression could provide information about the transcription profile in the presence of the drug, which can be used as an early indicator of the potential for toxic effects. For example, genes involved in cellular growth and controlled cell death, when disturbed, could lead to the growth of cancerous cells. Genome-wide studies can also help to find new genes involved in drug toxicity. Personal genome sequence information can be used to prescribe medications that will be most effective and least toxic on the basis of the individual patient's genotype. The gene signatures may not be completely accurate, but can be tested further before pathologic symptoms arise.

## Microbial Genomics: Metagenomics

Traditionally, microbiology has been taught with the view that microorganisms are best studied under **pure culture** conditions, which involves isolating a single type of cell and culturing it in the laboratory. Because microorganisms can go through several generations in a matter of hours, their gene expression profiles adapt to the new laboratory environment very quickly. In addition, the vast majority of bacterial species resist being cultured in isolation. Most microorganisms do not live as isolated entities, but in microbial communities known as biofilms. For all of these reasons, pure culture is not always the best way to study microorganisms. **Metagenomics** is the study of the collective genomes of multiple species that grow and interact in an environmental niche.



Metagenomics can be used to identify new species more rapidly and to analyze the effect of pollutants on the environment ([link](#)).



Metagenomics involves isolating DNA from multiple species within an environmental niche.

## Microbial Genomics: Creation of New Biofuels

Knowledge of the genomics of microorganisms is being used to find better ways to harness biofuels from algae and cyanobacteria. The primary sources of fuel today are coal, oil, wood, and other plant products, such as ethanol. Although plants are renewable resources, there is still a need to find more alternative renewable sources of energy to meet our population's energy demands. The microbial world is one of the largest resources for genes that encode new enzymes and produce new organic compounds, and it remains largely untapped. Microorganisms are used to create products, such as enzymes that are used in research, antibiotics, and other anti-microbial mechanisms. Microbial genomics is helping to develop diagnostic

tools, improved vaccines, new disease treatments, and advanced environmental cleanup techniques.

## **Mitochondrial Genomics**

Mitochondria are intracellular organelles that contain their own DNA. Mitochondrial DNA mutates at a rapid rate and is often used to study evolutionary relationships. Another feature that makes studying the mitochondrial genome interesting is that the mitochondrial DNA in most multicellular organisms is passed on from the mother during the process of fertilization. For this reason, mitochondrial genomics is often used to trace genealogy.

Information and clues obtained from DNA samples found at crime scenes have been used as evidence in court cases, and genetic markers have been used in forensic analysis. Genomic analysis has also become useful in this field. In 2001, the first use of genomics in forensics was published. It was a collaborative attempt between academic research institutions and the FBI to solve the mysterious cases of anthrax communicated via the US Postal Service. Using microbial genomics, researchers determined that a specific strain of anthrax was used in all the mailings.

## **Genomics in Agriculture**

Genomics can reduce the trials and failures involved in scientific research to a certain extent, which could improve the quality and quantity of crop yields in agriculture. Linking traits to genes or gene signatures helps to improve crop breeding to generate hybrids with the most desirable qualities. Scientists use genomic data to identify desirable traits, and then transfer those traits to a different organism. Scientists are discovering how genomics can improve the quality and quantity of agricultural production. For example, scientists could use desirable traits to create a useful product or enhance an existing product, such as making a drought-sensitive crop more tolerant of the dry season.

## **Section Summary**

Imagination is the only barrier to the applicability of genomics. Genomics is being applied to most fields of biology; it is being used for personalized medicine, prediction of disease risks at an individual level, the study of drug interactions before the conduct of clinical trials, and the study of microorganisms in the environment as opposed to the laboratory. It is also being applied to developments such as the generation of new biofuels, genealogical assessment using mitochondria, advances in forensic science, and improvements in agriculture.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) In 2011, the United States Preventative Services Task Force recommended against using the PSA test to screen healthy men for prostate cancer. Their recommendation is based on evidence that screening does not reduce the risk of death from prostate cancer. Prostate cancer often develops very slowly and does not cause problems, while the cancer treatment can have severe side effects. The *PCA3* test is considered to be more accurate, but screening may still result in men who would not have been harmed by the cancer itself suffering side effects from treatment. What do you think? Should all healthy men be screened for prostate cancer using the *PCA3* or PSA test? Should people in general be screened to find out if they have a genetic risk for cancer or other diseases?

---

#### Solution:

[\[link\]](#) There are no right or wrong answers to these questions. While it is true that prostate cancer treatment itself can be harmful, many men would rather be aware that they have cancer so they can monitor the disease and begin treatment if it progresses. And while genetic screening may be useful, it is expensive and may cause needless worry. People with certain risk factors may never develop the disease, and preventative treatments may do more harm than good.

## Review Questions

### Exercise:

**Problem:** Genomics can be used in agriculture to:

- a. generate new hybrid strains
- b. improve disease resistance
- c. improve yield
- d. all of the above

---

**Solution:**

D

### Exercise:

**Problem:** Genomics can be used on a personal level to:

- a. decrease transplant rejection
- b. Predict genetic diseases that a person may have inherited
- c. Determine the risks of genetic diseases for an individual's children
- d. All the above

---

**Solution:**

A

## Free Response

### Exercise:

**Problem:**

Explain why metagenomics is probably the most revolutionary application of genomics.

---

**Solution:**

Metagenomics is revolutionary because it replaced the practice of using pure cultures. Pure cultures were used to study individual species in the laboratory, but did not accurately represent what happens in the environment. Metagenomics studies the genomes of bacterial populations in their environmental niche.

**Exercise:****Problem:**

How can genomics be used to predict disease risk and treatment options?

---

**Solution:**

Genomics can provide the unique DNA sequence of an individual, which can be used for personalized medicine and treatment options.

**Glossary**

metagenomics

study of the collective genomes of multiple species that grow and interact in an environmental niche

pharmacogenomics

study of drug interactions with the genome or proteome; also called toxicogenomics

polygenic

phenotypic characteristic caused by two or more genes

pure culture

growth of a single type of cell in the laboratory

## Genomics and Proteomics

By the end of this section, you will be able to:

- Explain systems biology
- Describe a proteome
- Define protein signature

Proteins are the final products of genes, which help perform the function encoded by the gene. Proteins are composed of amino acids and play important roles in the cell. All enzymes (except ribozymes) are proteins that act as catalysts to affect the rate of reactions. Proteins are also regulatory molecules, and some are hormones. Transport proteins, such as hemoglobin, help transport oxygen to various organs. Antibodies that defend against foreign particles are also proteins. In the diseased state, protein function can be impaired because of changes at the genetic level or because of direct impact on a specific protein.

A **proteome** is the entire set of proteins produced by a cell type. Proteomes can be studied using the knowledge of genomes because genes code for mRNAs, and the mRNAs encode proteins. Although mRNA analysis is a step in the right direction, not all mRNAs are translated into proteins. The study of the function of proteomes is called **proteomics**. Proteomics complements genomics and is useful when scientists want to test their hypotheses that were based on genes. Even though all cells of a multicellular organism have the same set of genes, the set of proteins produced in different tissues is different and dependent on gene expression. Thus, the genome is constant, but the proteome varies and is dynamic within an organism. In addition, RNAs can be alternately spliced (cut and pasted to create novel combinations and novel proteins) and many proteins are modified after translation by processes such as proteolytic cleavage, phosphorylation, glycosylation, and ubiquitination. There are also protein-protein interactions, which complicate the study of proteomes. Although the genome provides a blueprint, the final architecture depends on several factors that can change the progression of events that generate the proteome.

Metabolomics is related to genomics and proteomics. **Metabolomics** involves the study of small molecule metabolites found in an organism. The

**metabolome** is the complete set of metabolites that are related to the genetic makeup of an organism. Metabolomics offers an opportunity to compare genetic makeup and physical characteristics, as well as genetic makeup and environmental factors. The goal of metabolome research is to identify, quantify, and catalogue all of the metabolites that are found in the tissues and fluids of living organisms.

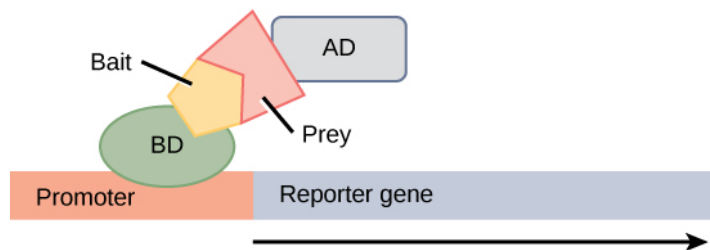
## Basic Techniques in Protein Analysis

The ultimate goal of proteomics is to identify or compare the proteins expressed from a given genome under specific conditions, study the interactions between the proteins, and use the information to predict cell behavior or develop drug targets. Just as the genome is analyzed using the basic technique of DNA sequencing, proteomics requires techniques for protein analysis. The basic technique for protein analysis, analogous to DNA sequencing, is mass spectrometry. Mass spectrometry is used to identify and determine the characteristics of a molecule. Advances in spectrometry have allowed researchers to analyze very small samples of protein. X-ray crystallography, for example, enables scientists to determine the three-dimensional structure of a protein crystal at atomic resolution. Another protein imaging technique, nuclear magnetic resonance (NMR), uses the magnetic properties of atoms to determine the three-dimensional structure of proteins in aqueous solution. Protein microarrays have also been used to study interactions between proteins. Large-scale adaptations of the basic two-hybrid screen ([\[link\]](#)) have provided the basis for protein microarrays. Computer software is used to analyze the vast amount of data generated for proteomic analysis.

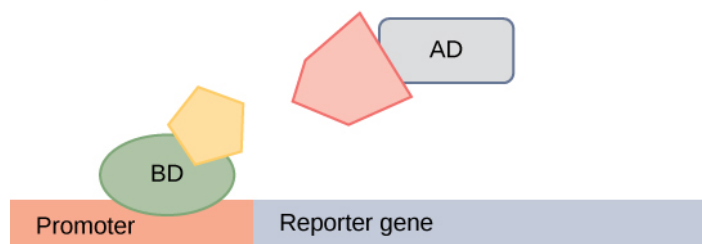
Genomic- and proteomic-scale analyses are part of systems biology. **Systems biology** is the study of whole biological systems (genomes and proteomes) based on interactions within the system. The European Bioinformatics Institute and the Human Proteome Organization (HUPO) are developing and establishing effective tools to sort through the enormous pile of systems biology data. Because proteins are the direct products of genes and reflect activity at the genomic level, it is natural to use proteomes to compare the protein profiles of different cells to identify proteins and genes involved in disease processes. Most pharmaceutical drug trials target



proteins. Information obtained from proteomics is being used to identify novel drugs and understand their mechanisms of action.



If the bait protein interacts with the prey protein, the transcription factor's activator domain binds to the binding domain, and transcription occurs.



If the prey doesn't catch the bait, no transcription occurs.

Two-hybrid screening is used to determine whether two proteins interact. In this method, a transcription factor is split into a DNA-binding domain (BD) and an activator domain (AD). The binding domain is able to bind the promoter in the absence of the activator domain, but it does not turn on transcription. A protein called the bait is attached to the BD, and a protein called the prey is attached to the AD. Transcription occurs only if the prey “catches” the bait.

The challenge of techniques used for proteomic analyses is the difficulty in detecting small quantities of proteins. Although mass spectrometry is good for detecting small amounts of proteins, variations in protein expression in diseased states can be difficult to discern. Proteins are naturally unstable molecules, which makes proteomic analysis much more difficult than genomic analysis.

## Cancer Proteomics

Genomes and proteomes of patients suffering from specific diseases are being studied to understand the genetic basis of the disease. The most prominent disease being studied with proteomic approaches is cancer. Proteomic approaches are being used to improve screening and early detection of cancer; this is achieved by identifying proteins whose expression is affected by the disease process. An individual protein is called a **biomarker**, whereas a set of proteins with altered expression levels is called a **protein signature**. For a biomarker or protein signature to be useful as a candidate for early screening and detection of a cancer, it must be secreted in body fluids, such as sweat, blood, or urine, such that large-scale screenings can be performed in a non-invasive fashion. The current problem with using biomarkers for the early detection of cancer is the high rate of false-negative results. A **false negative** is an incorrect test result that should have been positive. In other words, many cases of cancer go undetected, which makes biomarkers unreliable. Some examples of protein biomarkers used in cancer detection are CA-125 for ovarian cancer and PSA for prostate cancer. Protein signatures may be more reliable than biomarkers to detect cancer cells. Proteomics is also being used to develop individualized treatment plans, which involves the prediction of whether or not an individual will respond to specific drugs and the side effects that the individual may experience. Proteomics is also being used to predict the possibility of disease recurrence.

The National Cancer Institute has developed programs to improve the detection and treatment of cancer. The Clinical Proteomic Technologies for Cancer and the Early Detection Research Network are efforts to identify protein signatures specific to different types of cancers. The Biomedical

Proteomics Program is designed to identify protein signatures and design effective therapies for cancer patients.

## Section Summary

Proteomics is the study of the entire set of proteins expressed by a given type of cell under certain environmental conditions. In a multicellular organism, different cell types will have different proteomes, and these will vary with changes in the environment. Unlike a genome, a proteome is dynamic and in constant flux, which makes it both more complicated and more useful than the knowledge of genomes alone.

Proteomics approaches rely on protein analysis; these techniques are constantly being upgraded. Proteomics has been used to study different types of cancer. Different biomarkers and protein signatures are being used to analyze each type of cancer. The future goal is to have a personalized treatment plan for each individual.

## Review Questions

### Exercise:

**Problem:**What is a biomarker?

- a. the color coding of different genes
- b. a protein that is uniquely produced in a diseased state
- c. a molecule in the genome or proteome
- d. a marker that is genetically inherited

---

**Solution:**

B

### Exercise:

**Problem:**A protein signature is:

- a. the path followed by a protein after it is synthesized in the nucleus
- b. the path followed by a protein in the cytoplasm
- c. a protein expressed on the cell surface
- d. a unique set of proteins present in a diseased state

---

**Solution:**

D

## Free Response

**Exercise:**

**Problem:**

How has proteomics been used in cancer detection and treatment?

---

**Solution:**

Proteomics has provided a way to detect biomarkers and protein signatures, which have been used to screen for the early detection of cancer.

**Exercise:**

**Problem:**What is personalized medicine?

---

**Solution:**

Personalized medicine is the use of an individual's genomic sequence to predict the risk for specific diseases. When a disease does occur, it can be used to develop a personalized treatment plan.

## Glossary

biomarker

individual protein that is uniquely produced in a diseased state

false negative

incorrect test result that should have been positive

metabolome

complete set of metabolites which are related to the genetic makeup of an organism

metabolomics

study of small molecule metabolites found in an organism

protein signature

set of uniquely expressed proteins in the diseased state

proteome

entire set of proteins produced by a cell type

proteomics

study of the function of proteomes

systems biology

study of whole biological systems (genomes and proteomes) based on interactions within the system

## Introduction

class="introduction"

All  
organisms  
are products  
of evolution  
adapted to  
their  
environment.

(a) Saguaro  
(*Carnegiea  
gigantea*)  
can soak up  
750 liters of  
water in a  
single rain  
storm,  
enabling  
these cacti to  
survive the  
dry  
conditions of  
the Sonora  
desert in  
Mexico and  
the  
Southwestern  
United  
States. (b)  
The Andean  
semiaquatic  
lizard  
(*Potamites  
montanicola*)  
discovered in  
Peru in 2010

lives  
between  
1,570 to  
2,100 meters  
in elevation,  
and, unlike  
most lizards,  
is nocturnal  
and swims.

Scientists  
still do not  
know how  
these cold-  
blood  
animals are  
able to move  
in the cold  
(10 to 15°C)  
temperatures  
of the  
Andean  
night. (credit

a:  
modification  
of work by  
Gentry  
George, U.S.

Fish and  
Wildlife  
Service;  
credit b:  
modification  
of work by  
Germán  
Chávez and  
Diego

Vásquez,  
*ZooKeys*)



All species of living organisms, from bacteria to baboons to blueberries, evolved at some point from a different species. Although it may seem that living things today stay much the same, that is not the case—evolution is an ongoing process.

The theory of evolution is the unifying theory of biology, meaning it is the framework within which biologists ask questions about the living world. Its power is that it provides direction for predictions about living things that are borne out in experiment after experiment. The Ukrainian-born American geneticist Theodosius Dobzhansky famously wrote that “nothing makes sense in biology except in the light of evolution.”<sup>[[footnote](#)]</sup> He meant that the tenet that all life has evolved and diversified from a common ancestor is the foundation from which we approach all questions in biology. Theodosius Dobzhansky. “Biology, Molecular and Organismic.” *American Zoologist* 4, no. 4 (1964): 449.



## Understanding Evolution

By the end of this section, you will be able to:

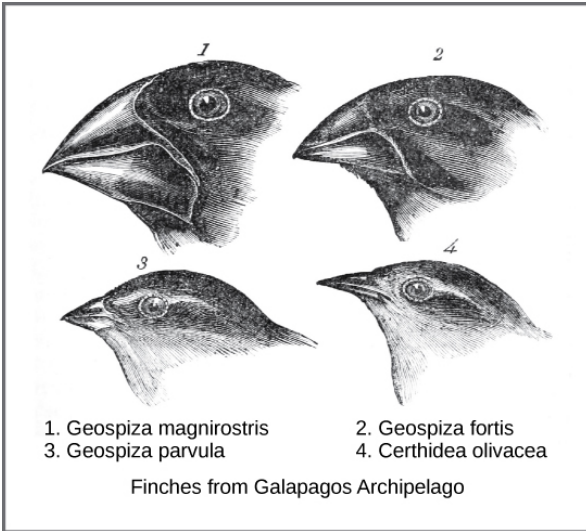
- Describe how the present-day theory of evolution was developed
- Define adaptation
- Explain convergent and divergent evolution
- Describe homologous and vestigial structures
- Discuss misconceptions about the theory of evolution

Evolution by natural selection describes a mechanism for how species change over time. That species change had been suggested and debated well before Darwin began to explore this idea. The view that species were static and unchanging was grounded in the writings of Plato, yet there were also ancient Greeks who expressed evolutionary ideas. In the eighteenth century, ideas about the evolution of animals were reintroduced by the naturalist Georges-Louis Leclerc Comte de Buffon who observed that various geographic regions have different plant and animal populations, even when the environments are similar. It was also accepted that there were extinct species.

During this time, James Hutton, a Scottish naturalist, proposed that geological change occurred gradually by the accumulation of small changes from processes operating like they are today over long periods of time. This contrasted with the predominant view that the geology of the planet was a consequence of catastrophic events occurring during a relatively brief past. Hutton's view was popularized in the nineteenth century by the geologist Charles Lyell who became a friend to Darwin. Lyell's ideas were influential on Darwin's thinking: Lyell's notion of the greater age of Earth gave more time for gradual change in species, and the process of change provided an analogy for gradual change in species. In the early nineteenth century, Jean-Baptiste Lamarck published a book that detailed a mechanism for evolutionary change. This mechanism is now referred to as an inheritance of acquired characteristics by which modifications in an individual are caused by its environment, or the use or disuse of a structure during its lifetime, could be inherited by its offspring and thus bring about change in a species. While this mechanism for evolutionary change was discredited, Lamarck's ideas were an important influence on evolutionary thought.

## Charles Darwin and Natural Selection

In the mid-nineteenth century, the actual mechanism for evolution was independently conceived of and described by two naturalists: Charles Darwin and Alfred Russel Wallace. Importantly, each naturalist spent time exploring the natural world on expeditions to the tropics. From 1831 to 1836, Darwin traveled around the world on *H.M.S. Beagle*, including stops in South America, Australia, and the southern tip of Africa. Wallace traveled to Brazil to collect insects in the Amazon rainforest from 1848 to 1852 and to the Malay Archipelago from 1854 to 1862. Darwin's journey, like Wallace's later journeys to the Malay Archipelago, included stops at several island chains, the last being the Galápagos Islands west of Ecuador. On these islands, Darwin observed species of organisms on different islands that were clearly similar, yet had distinct differences. For example, the ground finches inhabiting the Galápagos Islands comprised several species with a unique beak shape ([\[link\]](#)). The species on the islands had a graded series of beak sizes and shapes with very small differences between the most similar. He observed that these finches closely resembled another finch species on the mainland of South America. Darwin imagined that the island species might be species modified from one of the original mainland species. Upon further study, he realized that the varied beaks of each finch helped the birds acquire a specific type of food. For example, seed-eating finches had stronger, thicker beaks for breaking seeds, and insect-eating finches had spear-like beaks for stabbing their prey.



Darwin observed that beak shape varies among finch species. He postulated that the beak of an ancestral species had adapted over time to equip the finches to acquire different food sources.

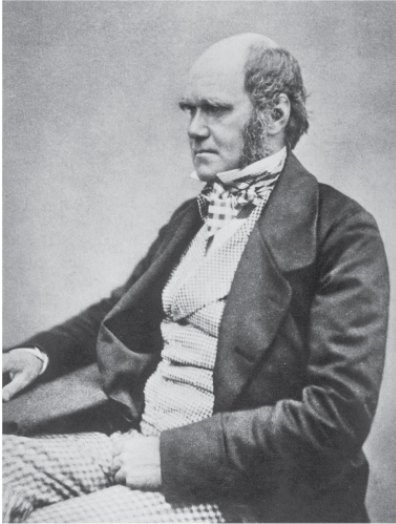
Wallace and Darwin both observed similar patterns in other organisms and they independently developed the same explanation for how and why such changes could take place. Darwin called this mechanism natural selection. **Natural selection**, also known as “survival of the fittest,” is the more prolific reproduction of individuals with favorable traits that survive environmental change because of those traits; this leads to evolutionary change.

For example, a population of giant tortoises found in the Galapagos Archipelago was observed by Darwin to have longer necks than those that lived on other islands with dry lowlands. These tortoises were “selected” because they could reach more leaves and access more food than those with short necks. In times of drought when fewer leaves would be available, those that could reach more leaves had a better chance to eat and survive than those that couldn’t reach the food source. Consequently, long-necked

tortoises would be more likely to be reproductively successful and pass the long-necked trait to their offspring. Over time, only long-necked tortoises would be present in the population.

Natural selection, Darwin argued, was an inevitable outcome of three principles that operated in nature. First, most characteristics of organisms are inherited, or passed from parent to offspring. Although no one, including Darwin and Wallace, knew how this happened at the time, it was a common understanding. Second, more offspring are produced than are able to survive, so resources for survival and reproduction are limited. The capacity for reproduction in all organisms outstrips the availability of resources to support their numbers. Thus, there is competition for those resources in each generation. Both Darwin and Wallace's understanding of this principle came from reading an essay by the economist Thomas Malthus who discussed this principle in relation to human populations. Third, offspring vary among each other in regard to their characteristics and those variations are inherited. Darwin and Wallace reasoned that offspring with inherited characteristics which allow them to best compete for limited resources will survive and have more offspring than those individuals with variations that are less able to compete. Because characteristics are inherited, these traits will be better represented in the next generation. This will lead to change in populations over generations in a process that Darwin called descent with modification. Ultimately, natural selection leads to greater adaptation of the population to its local environment; it is the only mechanism known for adaptive evolution.

Papers by Darwin and Wallace ([\[link\]](#)) presenting the idea of natural selection were read together in 1858 before the Linnean Society in London. The following year Darwin's book, *On the Origin of Species*, was published. His book outlined in considerable detail his arguments for evolution by natural selection.



(a)



(b)

Both (a) Charles Darwin and (b) Alfred Wallace wrote scientific papers on natural selection that were presented together before the Linnean Society in 1858.

Demonstrations of evolution by natural selection are time consuming and difficult to obtain. One of the best examples has been demonstrated in the very birds that helped to inspire Darwin's theory: the Galápagos finches. Peter and Rosemary Grant and their colleagues have studied Galápagos finch populations every year since 1976 and have provided important demonstrations of natural selection. The Grants found changes from one generation to the next in the distribution of beak shapes with the medium ground finch on the Galápagos island of Daphne Major. The birds have inherited variation in the bill shape with some birds having wide deep bills and others having thinner bills. During a period in which rainfall was higher than normal because of an El Niño, the large hard seeds that large-billed birds ate were reduced in number; however, there was an abundance of the small soft seeds which the small-billed birds ate. Therefore, survival and reproduction were much better in the following years for the small-billed birds. In the years following this El Niño, the Grants measured beak sizes in the population and found that the average bill size was smaller. Since bill size is an inherited trait, parents with smaller bills had more offspring and

the size of bills had evolved to be smaller. As conditions improved in 1987 and larger seeds became more available, the trend toward smaller average bill size ceased.

**Note:**

**Career Connection**

**Field Biologist**

Many people hike, explore caves, scuba dive, or climb mountains for recreation. People often participate in these activities hoping to see wildlife. Experiencing the outdoors can be incredibly enjoyable and invigorating. What if your job was to be outside in the wilderness? Field biologists by definition work outdoors in the “field.” The term field in this case refers to any location outdoors, even under water. A field biologist typically focuses research on a certain species, group of organisms, or a single habitat ([\[link\]](#)).



A field biologist tranquilizes a polar bear for study. (credit: Karen Rhode)

One objective of many field biologists includes discovering new species that have never been recorded. Not only do such findings expand our understanding of the natural world, but they also lead to important innovations in fields such as medicine and agriculture. Plant and microbial

species, in particular, can reveal new medicinal and nutritive knowledge. Other organisms can play key roles in ecosystems or be considered rare and in need of protection. When discovered, these important species can be used as evidence for environmental regulations and laws.

## Processes and Patterns of Evolution

Natural selection can only take place if there is **variation**, or differences, among individuals in a population. Importantly, these differences must have some genetic basis; otherwise, the selection will not lead to change in the next generation. This is critical because variation among individuals can be caused by non-genetic reasons such as an individual being taller because of better nutrition rather than different genes.

Genetic diversity in a population comes from two main mechanisms: mutation and sexual reproduction. Mutation, a change in DNA, is the ultimate source of new alleles, or new genetic variation in any population. The genetic changes caused by mutation can have one of three outcomes on the phenotype. A mutation affects the phenotype of the organism in a way that gives it reduced fitness—lower likelihood of survival or fewer offspring. A mutation may produce a phenotype with a beneficial effect on fitness. And, many mutations will also have no effect on the fitness of the phenotype; these are called neutral mutations. Mutations may also have a whole range of effect sizes on the fitness of the organism that expresses them in their phenotype, from a small effect to a great effect. Sexual reproduction also leads to genetic diversity: when two parents reproduce, unique combinations of alleles assemble to produce the unique genotypes and thus phenotypes in each of the offspring.

A heritable trait that helps the survival and reproduction of an organism in its present environment is called an **adaptation**. Scientists describe groups of organisms becoming adapted to their environment when a change in the range of genetic variation occurs over time that increases or maintains the “fit” of the population to its environment. The webbed feet of platypuses are an adaptation for swimming. The snow leopards’ thick fur is an



adaptation for living in the cold. The cheetahs' fast speed is an adaptation for catching prey.

Whether or not a trait is favorable depends on the environmental conditions at the time. The same traits are not always selected because environmental conditions can change. For example, consider a species of plant that grew in a moist climate and did not need to conserve water. Large leaves were selected because they allowed the plant to obtain more energy from the sun. Large leaves require more water to maintain than small leaves, and the moist environment provided favorable conditions to support large leaves. After thousands of years, the climate changed, and the area no longer had excess water. The direction of natural selection shifted so that plants with small leaves were selected because those populations were able to conserve water to survive the new environmental conditions.

The evolution of species has resulted in enormous variation in form and function. Sometimes, evolution gives rise to groups of organisms that become tremendously different from each other. When two species evolve in diverse directions from a common point, it is called **divergent evolution**. Such divergent evolution can be seen in the forms of the reproductive organs of flowering plants which share the same basic anatomies; however, they can look very different as a result of selection in different physical environments and adaptation to different kinds of pollinators ([\[link\]](#)).



Flowering plants evolved from a common ancestor. Notice that the (a) dense blazing



star (*Liatrus spicata*) and the (b) purple coneflower (*Echinacea purpurea*) vary in appearance, yet both share a similar basic morphology. (credit a: modification of work by Drew Avery; credit b: modification of work by Cory Zanker)

In other cases, similar phenotypes evolve independently in distantly related species. For example, flight has evolved in both bats and insects, and they both have structures we refer to as wings, which are adaptations to flight. However, the wings of bats and insects have evolved from very different original structures. This phenomenon is called **convergent evolution**, where similar traits evolve independently in species that do not share a common ancestry. The two species came to the same function, flying, but did so separately from each other.

These physical changes occur over enormous spans of time and help explain how evolution occurs. Natural selection acts on individual organisms, which in turn can shape an entire species. Although natural selection may work in a single generation on an individual, it can take thousands or even millions of years for the genotype of an entire species to evolve. It is over these large time spans that life on earth has changed and continues to change.

## Evidence of Evolution

The evidence for evolution is compelling and extensive. Looking at every level of organization in living systems, biologists see the signature of past and present evolution. Darwin dedicated a large portion of his book, *On the Origin of Species*, to identifying patterns in nature that were consistent with evolution, and since Darwin, our understanding has become clearer and broader.

## Fossils

Fossils provide solid evidence that organisms from the past are not the same as those found today, and fossils show a progression of evolution. Scientists determine the age of fossils and categorize them from all over the world to determine when the organisms lived relative to each other. The resulting fossil record tells the story of the past and shows the evolution of form over millions of years ([\[link\]](#)). For example, scientists have recovered highly detailed records showing the evolution of humans and horses ([\[link\]](#)). The whale flipper shares a similar morphology to appendages of birds and mammals ([\[link\]](#)) indicating that these species share a common ancestor.



(a)

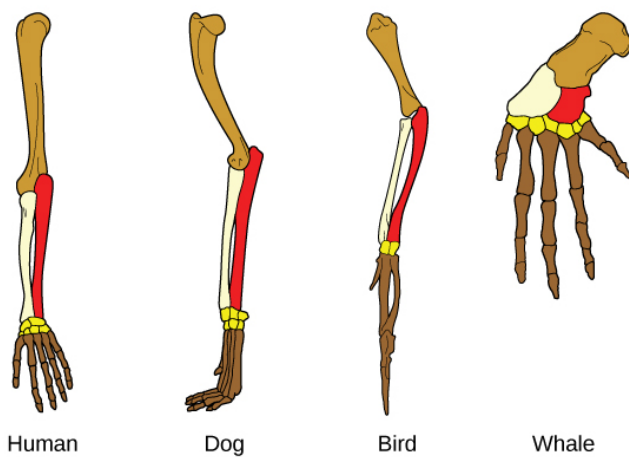


(b)

In this (a) display, fossil hominids are arranged from oldest (bottom) to newest (top). As hominids evolved, the shape of the skull changed. An artist's rendition of (b) extinct species of the genus *Equus* reveals that these ancient species resembled the modern horse (*Equus ferus*) but varied in size.

## Anatomy and Embryology

Another type of evidence for evolution is the presence of structures in organisms that share the same basic form. For example, the bones in the appendages of a human, dog, bird, and whale all share the same overall construction ([link](#)) resulting from their origin in the appendages of a common ancestor. Over time, evolution led to changes in the shapes and sizes of these bones in different species, but they have maintained the same overall layout. Scientists call these synonymous parts **homologous structures**.



The similar construction of these appendages indicates that these organisms share a common ancestor.

Some structures exist in organisms that have no apparent function at all, and appear to be residual parts from a past common ancestor. These unused structures without function are called **vestigial structures**. Other examples of vestigial structures are wings on flightless birds, leaves on some cacti, and hind leg bones in whales.

**Note:**

## Link to Learning



Visit this [interactive site](#) to guess which bones structures are homologous and which are analogous, and see examples of evolutionary adaptations to illustrate these concepts.

Another evidence of evolution is the convergence of form in organisms that share similar environments. For example, species of unrelated animals, such as the arctic fox and ptarmigan, living in the arctic region have been selected for seasonal white phenotypes during winter to blend with the snow and ice ([link](#)ab). These similarities occur not because of common ancestry, but because of similar selection pressures—the benefits of not being seen by predators.



(a)

(b)

The white winter coat of the (a) arctic fox

and the (b) ptarmigan's plumage are adaptations to their environments. (credit a: modification of work by Keith Morehouse)

Embryology, the study of the development of the anatomy of an organism to its adult form, also provides evidence of relatedness between now widely divergent groups of organisms. Mutational tweaking in the embryo can have such magnified consequences in the adult that embryo formation tends to be conserved. As a result, structures that are absent in some groups often appear in their embryonic forms and disappear by the time the adult or juvenile form is reached. For example, all vertebrate embryos, including humans, exhibit gill slits and tails at some point in their early development. These disappear in the adults of terrestrial groups but are maintained in adult forms of aquatic groups such as fish and some amphibians. Great ape embryos, including humans, have a tail structure during their development that is lost by the time of birth.

## **Biogeography**

The geographic distribution of organisms on the planet follows patterns that are best explained by evolution in conjunction with the movement of tectonic plates over geological time. Broad groups that evolved before the breakup of the supercontinent Pangaea (about 200 million years ago) are distributed worldwide. Groups that evolved since the breakup appear uniquely in regions of the planet, such as the unique flora and fauna of northern continents that formed from the supercontinent Laurasia and of the southern continents that formed from the supercontinent Gondwana. The presence of members of the plant family Proteaceae in Australia, southern Africa, and South America is best by their presence prior to the southern supercontinent Gondwana breaking up.

The great diversification of marsupials in Australia and the absence of other mammals reflect Australia's long isolation. Australia has an abundance of endemic species—species found nowhere else—which is typical of islands

whose isolation by expanses of water prevents species to migrate. Over time, these species diverge evolutionarily into new species that look very different from their ancestors that may exist on the mainland. The marsupials of Australia, the finches on the Galápagos, and many species on the Hawaiian Islands are all unique to their one point of origin, yet they display distant relationships to ancestral species on mainlands.

## **Molecular Biology**

Like anatomical structures, the structures of the molecules of life reflect descent with modification. Evidence of a common ancestor for all of life is reflected in the universality of DNA as the genetic material and in the near universality of the genetic code and the machinery of DNA replication and expression. Fundamental divisions in life between the three domains are reflected in major structural differences in otherwise conservative structures such as the components of ribosomes and the structures of membranes. In general, the relatedness of groups of organisms is reflected in the similarity of their DNA sequences—exactly the pattern that would be expected from descent and diversification from a common ancestor.

DNA sequences have also shed light on some of the mechanisms of evolution. For example, it is clear that the evolution of new functions for proteins commonly occurs after gene duplication events that allow the free modification of one copy by mutation, selection, or drift (changes in a population's gene pool resulting from chance), while the second copy continues to produce a functional protein.

## **Misconceptions of Evolution**

Although the theory of evolution generated some controversy when it was first proposed, it was almost universally accepted by biologists, particularly younger biologists, within 20 years after publication of *On the Origin of Species*. Nevertheless, the theory of evolution is a difficult concept and misconceptions about how it works abound.

**Note:****Link to Learning**

This [site](#) addresses some of the main misconceptions associated with the theory of evolution.

## **Evolution Is Just a Theory**

Critics of the theory of evolution dismiss its importance by purposefully confounding the everyday usage of the word “theory” with the way scientists use the word. In science, a “theory” is understood to be a body of thoroughly tested and verified explanations for a set of observations of the natural world. Scientists have a theory of the atom, a theory of gravity, and the theory of relativity, each of which describes understood facts about the world. In the same way, the theory of evolution describes facts about the living world. As such, a theory in science has survived significant efforts to discredit it by scientists. In contrast, a “theory” in common vernacular is a word meaning a guess or suggested explanation; this meaning is more akin to the scientific concept of “hypothesis.” When critics of evolution say evolution is “just a theory,” they are implying that there is little evidence supporting it and that it is still in the process of being rigorously tested. This is a mischaracterization.

## **Individuals Evolve**

Evolution is the change in genetic composition of a population over time, specifically over generations, resulting from differential reproduction of individuals with certain alleles. Individuals do change over their lifetime,

obviously, but this is called development and involves changes programmed by the set of genes the individual acquired at birth in coordination with the individual's environment. When thinking about the evolution of a characteristic, it is probably best to think about the change of the average value of the characteristic in the population over time. For example, when natural selection leads to bill-size change in medium-ground finches in the Galápagos, this does not mean that individual bills on the finches are changing. If one measures the average bill size among all individuals in the population at one time and then measures the average bill size in the population several years later, this average value will be different as a result of evolution. Although some individuals may survive from the first time to the second, they will still have the same bill size; however, there will be many new individuals that contribute to the shift in average bill size.

## **Evolution Explains the Origin of Life**

It is a common misunderstanding that evolution includes an explanation of life's origins. Conversely, some of the theory's critics believe that it cannot explain the origin of life. The theory does not try to explain the origin of life. The theory of evolution explains how populations change over time and how life diversifies the origin of species. It does not shed light on the beginnings of life including the origins of the first cells, which is how life is defined. The mechanisms of the origin of life on Earth are a particularly difficult problem because it occurred a very long time ago, and presumably it just occurred once. Importantly, biologists believe that the presence of life on Earth precludes the possibility that the events that led to life on Earth can be repeated because the intermediate stages would immediately become food for existing living things.

However, once a mechanism of inheritance was in place in the form of a molecule like DNA either within a cell or pre-cell, these entities would be subject to the principle of natural selection. More effective reproducers would increase in frequency at the expense of inefficient reproducers. So while evolution does not explain the origin of life, it may have something to say about some of the processes operating once pre-living entities acquired certain properties.



## Organisms Evolve on Purpose

Statements such as “organisms evolve in response to a change in an environment” are quite common, but such statements can lead to two types of misunderstandings. First, the statement must not be understood to mean that individual organisms evolve. The statement is shorthand for “a population evolves in response to a changing environment.” However, a second misunderstanding may arise by interpreting the statement to mean that the evolution is somehow intentional. A changed environment results in some individuals in the population, those with particular phenotypes, benefiting and therefore producing proportionately more offspring than other phenotypes. This results in change in the population if the characteristics are genetically determined.

It is also important to understand that the variation that natural selection works on is already in a population and does not arise in response to an environmental change. For example, applying antibiotics to a population of bacteria will, over time, select a population of bacteria that are resistant to antibiotics. The resistance, which is caused by a gene, did not arise by mutation because of the application of the antibiotic. The gene for resistance was already present in the gene pool of the bacteria, likely at a low frequency. The antibiotic, which kills the bacterial cells without the resistance gene, strongly selects individuals that are resistant, since these would be the only ones that survived and divided. Experiments have demonstrated that mutations for antibiotic resistance do not arise as a result of antibiotic.

In a larger sense, evolution is not goal directed. Species do not become “better” over time; they simply track their changing environment with adaptations that maximize their reproduction in a particular environment at a particular time. Evolution has no goal of making faster, bigger, more complex, or even smarter species, despite the commonness of this kind of language in popular discourse. What characteristics evolve in a species are a function of the variation present and the environment, both of which are constantly changing in a non-directional way. What trait is fit in one environment at one time may well be fatal at some point in the future. This holds equally well for a species of insect as it does the human species.

## Section Summary

Evolution is the process of adaptation through mutation which allows more desirable characteristics to be passed to the next generation. Over time, organisms evolve more characteristics that are beneficial to their survival. For living organisms to adapt and change to environmental pressures, genetic variation must be present. With genetic variation, individuals have differences in form and function that allow some to survive certain conditions better than others. These organisms pass their favorable traits to their offspring. Eventually, environments change, and what was once a desirable, advantageous trait may become an undesirable trait and organisms may further evolve. Evolution may be convergent with similar traits evolving in multiple species or divergent with diverse traits evolving in multiple species that came from a common ancestor. Evidence of evolution can be observed by means of DNA code and the fossil record, and also by the existence of homologous and vestigial structures.

## Review Questions

### Exercise:

#### Problem:

Which scientific concept did Charles Darwin and Alfred Wallace independently discover?

- a. mutation
- b. natural selection
- c. overbreeding
- d. sexual reproduction

---

#### Solution:

B

### Exercise:

**Problem:**

Which of the following situations will lead to natural selection?

- a. The seeds of two plants land near each other and one grows larger than the other.
- b. Two types of fish eat the same kind of food, and one is better able to gather food than the other.
- c. Male lions compete for the right to mate with females, with only one possible winner.
- d. all of the above

---

**Solution:**

D

**Exercise:**

**Problem:** Which description is an example of a phenotype?

- a. A certain duck has a blue beak.
- b. A mutation occurred to a flower.
- c. Most cheetahs live solitary lives.
- d. both a and c

---

**Solution:**

D

**Exercise:****Problem:**

Which situation is most likely an example of convergent evolution?

- a. Squid and humans have eyes similar in structure.
- b. Worms and snakes both move without legs.

- c. Some bats and birds have wings that allow them to fly
- d. all of the above

---

**Solution:**

D

**Free Response**

**Exercise:**

**Problem:**

If a person scatters a handful of garden pea plant seeds in one area, how would natural selection work in this situation?

---

**Solution:**

The plants that can best use the resources of the area, including competing with other individuals for those resources will produce more seeds themselves and those traits that allowed them to better use the resources will increase in the population of the next generation.

**Exercise:**

**Problem:**

Why do scientists consider vestigial structures evidence for evolution?

---

**Solution:**

Vestigial structures are considered evidence for evolution because most structures do not exist in an organism without serving some function either presently or in the past. A vestigial structure indicates a past form or function that has since changed, but the structure remains present because it had a function in the ancestor.

**Exercise:**

**Problem:**

How does the scientific meaning of “theory” differ from the common vernacular meaning?

---

**Solution:**

In science, a theory is a thoroughly tested and verified set of explanations for a body of observations of nature. It is the strongest form of knowledge in science. In contrast, a theory in common vernacular can mean a guess or speculation about something, meaning that the knowledge implied by the theory is very weak.

**Exercise:****Problem:**

Explain why the statement that a monkey is more evolved than a mouse is incorrect.

---

**Solution:**

The statement implies that there is a goal to evolution and that the monkey represents greater progress to that goal than the mouse. Both species are likely to be well adapted to their particular environments, which is the outcome of natural selection.

**Glossary****adaptation**

heritable trait or behavior in an organism that aids in its survival and reproduction in its present environment

**convergent evolution**

process by which groups of organisms independently evolve to similar forms

**divergent evolution**

process by which groups of organisms evolve in diverse directions from a common point

homologous structures

parallel structures in diverse organisms that have a common ancestor

natural selection

reproduction of individuals with favorable genetic traits that survive environmental change because of those traits, leading to evolutionary change

variation

genetic differences among individuals in a population

vestigial structure

physical structure present in an organism but that has no apparent function and appears to be from a functional structure in a distant ancestor

## Formation of New Species

By the end of this section, you will be able to:

- Define species and describe how species are identified as different
- Describe genetic variables that lead to speciation
- Identify prezygotic and postzygotic reproductive barriers
- Explain allopatric and sympatric speciation
- Describe adaptive radiation

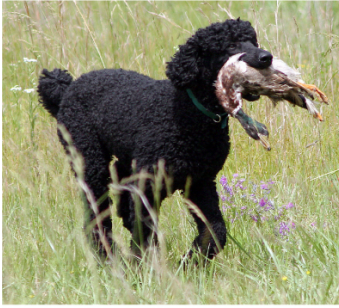
Although all life on earth shares various genetic similarities, only certain organisms combine genetic information by sexual reproduction and have offspring that can then successfully reproduce. Scientists call such organisms members of the same biological species.

## Species and the Ability to Reproduce

A **species** is a group of individual organisms that interbreed and produce fertile, viable offspring. According to this definition, one species is distinguished from another when, in nature, it is not possible for matings between individuals from each species to produce fertile offspring.

Members of the same species share both external and internal characteristics, which develop from their DNA. The closer relationship two organisms share, the more DNA they have in common, just like people and their families. People's DNA is likely to be more like their father or mother's DNA than their cousin or grandparent's DNA. Organisms of the same species have the highest level of DNA alignment and therefore share characteristics and behaviors that lead to successful reproduction.

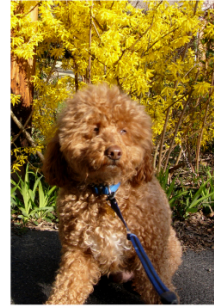
Species' appearance can be misleading in suggesting an ability or inability to mate. For example, even though domestic dogs (*Canis lupus familiaris*) display phenotypic differences, such as size, build, and coat, most dogs can interbreed and produce viable puppies that can mature and sexually reproduce ([link](#)).



(a)



(b)



(c)

The (a) poodle and (b) cocker spaniel can reproduce to produce a breed known as (c) the cockapoo. (credit a: modification of work by Sally Eller, Tom Reese; credit b: modification of work by Jeremy McWilliams; credit c: modification of work by Kathleen Conklin)

In other cases, individuals may appear similar although they are not members of the same species. For example, even though bald eagles (*Haliaeetus leucocephalus*) and African fish eagles (*Haliaeetus vocifer*) are both birds and eagles, each belongs to a separate species group ([\[link\]](#)). If humans were to artificially intervene and fertilize the egg of a bald eagle with the sperm of an African fish eagle and a chick did hatch, that offspring, called a **hybrid** (a cross between two species), would probably be infertile—unable to successfully reproduce after it reached maturity. Different species may have different genes that are active in development; therefore, it may not be possible to develop a viable offspring with two different sets of directions. Thus, even though hybridization may take place, the two species still remain separate.





(a)



(b)

The (a) African fish eagle is similar in appearance to the (b) bald eagle, but the two birds are members of different species.  
(credit a: modification of work by Nigel Wedge; credit b: modification of work by U.S. Fish and Wildlife Service)

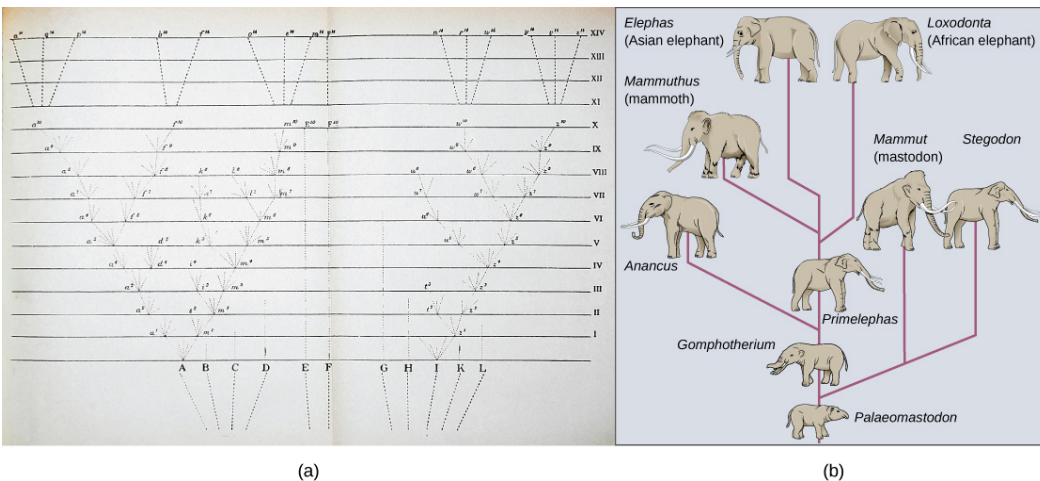
Populations of species share a gene pool: a collection of all the variants of genes in the species. Again, the basis to any changes in a group or population of organisms must be genetic for this is the only way to share and pass on traits. When variations occur within a species, they can only be passed to the next generation along two main pathways: asexual reproduction or sexual reproduction. The change will be passed on asexually simply if the reproducing cell possesses the changed trait. For the changed trait to be passed on by sexual reproduction, a gamete, such as a sperm or egg cell, must possess the changed trait. In other words, sexually-reproducing organisms can experience several genetic changes in their body cells, but if these changes do not occur in a sperm or egg cell, the changed trait will never reach the next generation. Only heritable traits can evolve. Therefore, reproduction plays a paramount role for genetic change to take root in a population or species. In short, organisms must be able to reproduce with each other to pass new traits to offspring.

## Speciation

The biological definition of species, which works for sexually reproducing organisms, is a group of actually or potentially interbreeding individuals.

There are exceptions to this rule. Many species are similar enough that hybrid offspring are possible and may often occur in nature, but for the majority of species this rule generally holds. In fact, the presence in nature of hybrids between similar species suggests that they may have descended from a single interbreeding species, and the speciation process may not yet be completed.

Given the extraordinary diversity of life on the planet there must be mechanisms for **speciation**: the formation of two species from one original species. Darwin envisioned this process as a branching event and diagrammed the process in the only illustration found in *On the Origin of Species* ([link](#)[a](#)). Compare this illustration to the diagram of elephant evolution ([link](#)[b](#)), which shows that as one species changes over time, it branches to form more than one new species, repeatedly, as long as the population survives or until the organism becomes extinct.



The only illustration in Darwin's *On the Origin of Species* is (a) a diagram showing speciation events leading to biological diversity. The diagram shows similarities to phylogenetic charts that are drawn today to illustrate the relationships of species. (b) Modern elephants evolved from the *Palaeomastodon*, a species that lived in Egypt 35–50 million years ago.

For speciation to occur, two new populations must be formed from one original population and they must evolve in such a way that it becomes impossible for individuals from the two new populations to interbreed. Biologists have proposed mechanisms by which this could occur that fall into two broad categories. **Allopatric speciation** (allo- = "other"; -patric = "homeland") involves geographic separation of populations from a parent species and subsequent evolution. **Sympatric speciation** (sym- = "same"; -patric = "homeland") involves speciation occurring within a parent species remaining in one location.

Biologists think of speciation events as the splitting of one ancestral species into two descendant species. There is no reason why there might not be more than two species formed at one time except that it is less likely and multiple events can be conceptualized as single splits occurring close in time.

## **Allopatric Speciation**

A geographically continuous population has a gene pool that is relatively homogeneous. Gene flow, the movement of alleles across the range of the species, is relatively free because individuals can move and then mate with individuals in their new location. Thus, the frequency of an allele at one end of a distribution will be similar to the frequency of the allele at the other end. When populations become geographically discontinuous, that free-flow of alleles is prevented. When that separation lasts for a period of time, the two populations are able to evolve along different trajectories. Thus, their allele frequencies at numerous genetic loci gradually become more and more different as new alleles independently arise by mutation in each population. Typically, environmental conditions, such as climate, resources, predators, and competitors for the two populations will differ causing natural selection to favor divergent adaptations in each group.

Isolation of populations leading to allopatric speciation can occur in a variety of ways: a river forming a new branch, erosion forming a new valley, a group of organisms traveling to a new location without the ability to return, or seeds floating over the ocean to an island. The nature of the

geographic separation necessary to isolate populations depends entirely on the biology of the organism and its potential for dispersal. If two flying insect populations took up residence in separate nearby valleys, chances are, individuals from each population would fly back and forth continuing gene flow. However, if two rodent populations became divided by the formation of a new lake, continued gene flow would be unlikely; therefore, speciation would be more likely.

Biologists group allopatric processes into two categories: dispersal and vicariance. **Dispersal** is when a few members of a species move to a new geographical area, and **vicariance** is when a natural situation arises to physically divide organisms.

Scientists have documented numerous cases of allopatric speciation taking place. For example, along the west coast of the United States, two separate sub-species of spotted owls exist. The northern spotted owl has genetic and phenotypic differences from its close relative: the Mexican spotted owl, which lives in the south ([link](#)).



The northern spotted owl and

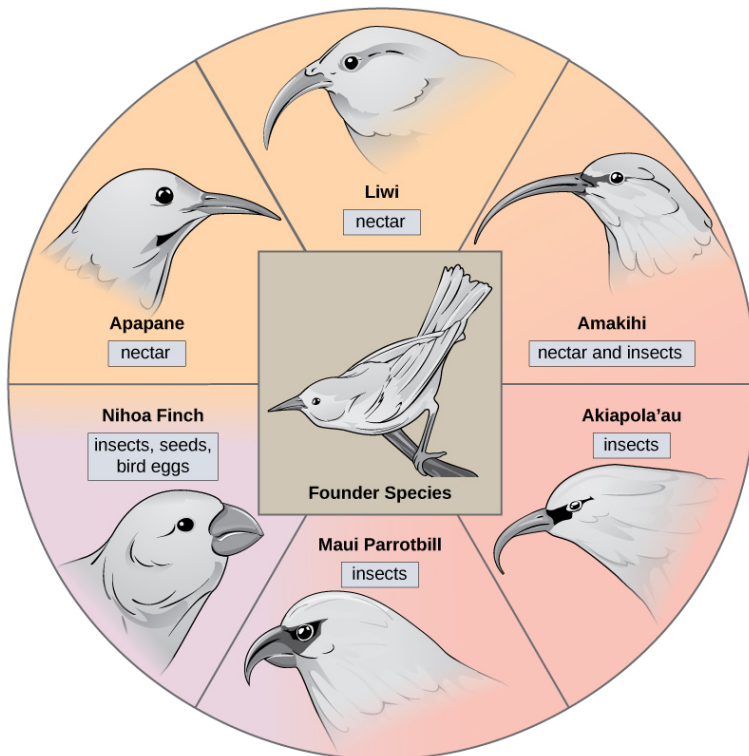
the Mexican spotted owl inhabit geographically separate locations with different climates and ecosystems. The owl is an example of allopatric speciation. (credit "northern spotted owl": modification of work by John and Karen Hollingsworth; credit "Mexican spotted owl": modification of work by Bill Radke)

Additionally, scientists have found that the further the distance between two groups that once were the same species, the more likely it is that speciation will occur. This seems logical because as the distance increases, the various environmental factors would likely have less in common than locations in close proximity. Consider the two owls: in the north, the climate is cooler than in the south; the types of organisms in each ecosystem differ, as do their behaviors and habits; also, the hunting habits and prey choices of the southern owls vary from the northern owls. These variances can lead to evolved differences in the owls, and speciation likely will occur.

## **Adaptive Radiation**

In some cases, a population of one species disperses throughout an area, and each finds a distinct niche or isolated habitat. Over time, the varied demands of their new lifestyles lead to multiple speciation events originating from a single species. This is called **adaptive radiation** because many adaptations evolve from a single point of origin; thus, causing the species to radiate into several new ones. Island archipelagos like the Hawaiian Islands provide an ideal context for adaptive radiation events because water surrounds each island which leads to geographical isolation

for many organisms. The Hawaiian honeycreeper illustrates one example of adaptive radiation. From a single species, called the founder species, numerous species have evolved, including the six shown in [\[link\]](#).



The honeycreeper birds illustrate adaptive radiation. From one original species of bird, multiple others evolved, each with its own distinctive characteristics.

Notice the differences in the species' beaks in [\[link\]](#). Evolution in response to natural selection based on specific food sources in each new habitat led to evolution of a different beak suited to the specific food source. The seed-eating bird has a thicker, stronger beak which is suited to break hard nuts. The nectar-eating birds have long beaks to dip into flowers to reach the nectar. The insect-eating birds have beaks like swords, appropriate for

stabbing and impaling insects. Darwin's finches are another example of adaptive radiation in an archipelago.

**Note:**

Link to Learning



Click through this [interactive site](#) to see how island birds evolved in evolutionary increments from 5 million years ago to today.

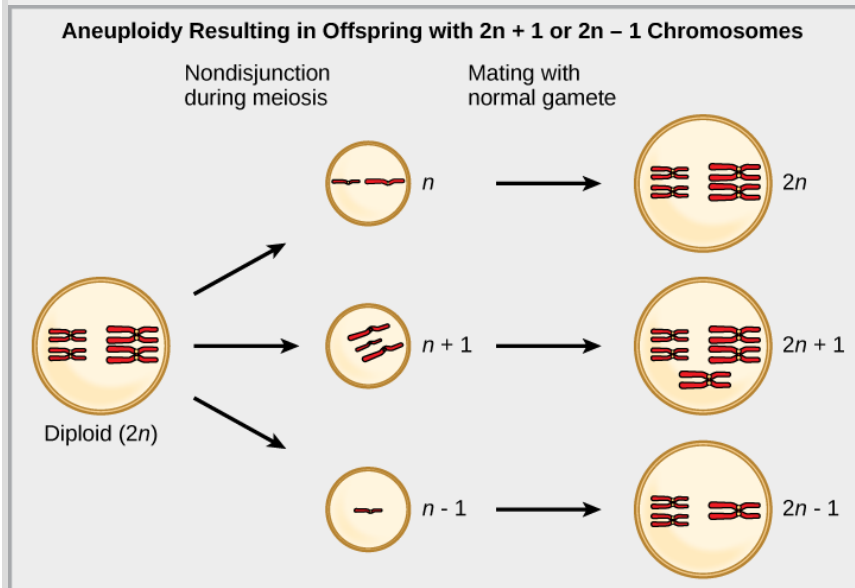
## Sympatric Speciation

Can divergence occur if no physical barriers are in place to separate individuals who continue to live and reproduce in the same habitat? The answer is yes. The process of speciation within the same space is called sympatric speciation; the prefix “sym” means same, so “sympatric” means “same homeland” in contrast to “allopatric” meaning “other homeland.” A number of mechanisms for sympatric speciation have been proposed and studied.

One form of sympatric speciation can begin with a serious chromosomal error during cell division. In a normal cell division event chromosomes replicate, pair up, and then separate so that each new cell has the same number of chromosomes. However, sometimes the pairs separate and the end cell product has too many or too few individual chromosomes in a condition called **aneuploidy** ([link](#)).

## Note:

### Art Connection



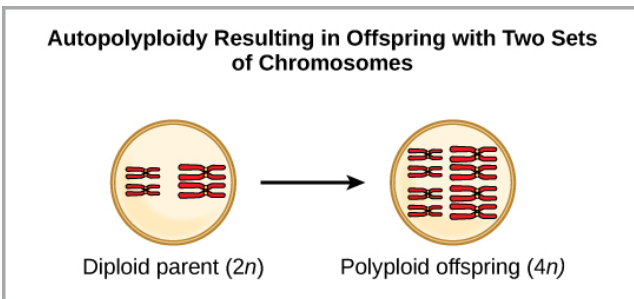
Aneuploidy results when the gametes have too many or too few chromosomes due to nondisjunction during meiosis. In the example shown here, the resulting offspring will have  $2n+1$  or  $2n-1$  chromosomes

Which is most likely to survive, offspring with  $2n+1$  chromosomes or offspring with  $2n-1$  chromosomes?

Polyploidy is a condition in which a cell or organism has an extra set, or sets, of chromosomes. Scientists have identified two main types of polyploidy that can lead to reproductive isolation of an individual in the polyploidy state. Reproductive isolation is the inability to interbreed. In some cases, a polyploid individual will have two or more complete sets of chromosomes from its own species in a condition called **autopolyploidy** ([link](#)). The prefix “auto-” means “self,” so the term means multiple chromosomes from one’s own species. Polyploidy results from an error in



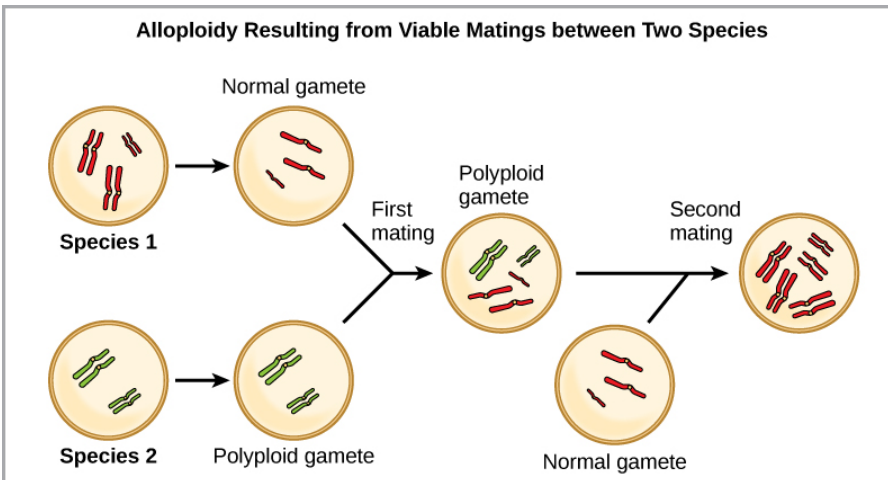
meiosis in which all of the chromosomes move into one cell instead of separating.



Autopolyploidy results when  
mitosis is not followed by  
cytokinesis.

For example, if a plant species with  $2n = 6$  produces autopolyploid gametes that are also diploid ( $2n = 6$ , when they should be  $n = 3$ ), the gametes now have twice as many chromosomes as they should have. These new gametes will be incompatible with the normal gametes produced by this plant species. However, they could either self-pollinate or reproduce with other autopolyploid plants with gametes having the same diploid number. In this way, sympatric speciation can occur quickly by forming offspring with  $4n$  called a tetraploid. These individuals would immediately be able to reproduce only with those of this new kind and not those of the ancestral species.

The other form of polyploidy occurs when individuals of two different species reproduce to form a viable offspring called an **allopolyploid**. The prefix “allo-” means “other” (recall from allopatric): therefore, an allopolyploid occurs when gametes from two different species combine. [\[link\]](#) illustrates one possible way an allopolyploid can form. Notice how it takes two generations, or two reproductive acts, before the viable fertile hybrid results.



Allopolyploidy results when two species mate to produce viable offspring. In the example shown, a normal gamete from one species fuses with a polyploidy gamete from another. Two matings are necessary to produce viable offspring.

The cultivated forms of wheat, cotton, and tobacco plants are all allopolyploids. Although polyploidy occurs occasionally in animals, it takes place most commonly in plants. (Animals with any of the types of chromosomal aberrations described here are unlikely to survive and produce normal offspring.) Scientists have discovered more than half of all plant species studied relate back to a species evolved through polyploidy. With such a high rate of polyploidy in plants, some scientists hypothesize that this mechanism takes place more as an adaptation than as an error.

## Reproductive Isolation

Given enough time, the genetic and phenotypic divergence between populations will affect characters that influence reproduction: if individuals of the two populations were to be brought together, mating would be less likely, but if mating occurred, offspring would be non-viable or infertile. Many types of diverging characters may affect the **reproductive isolation**, the ability to interbreed, of the two populations.

Reproductive isolation can take place in a variety of ways. Scientists organize them into two groups: prezygotic barriers and postzygotic barriers. Recall that a zygote is a fertilized egg: the first cell of the development of an organism that reproduces sexually. Therefore, a **prezygotic barrier** is a mechanism that blocks reproduction from taking place; this includes barriers that prevent fertilization when organisms attempt reproduction. A **postzygotic barrier** occurs after zygote formation; this includes organisms that don't survive the embryonic stage and those that are born sterile.

Some types of prezygotic barriers prevent reproduction entirely. Many organisms only reproduce at certain times of the year, often just annually. Differences in breeding schedules, called **temporal isolation**, can act as a form of reproductive isolation. For example, two species of frogs inhabit the same area, but one reproduces from January to March, whereas the other reproduces from March to May ([link](#)).



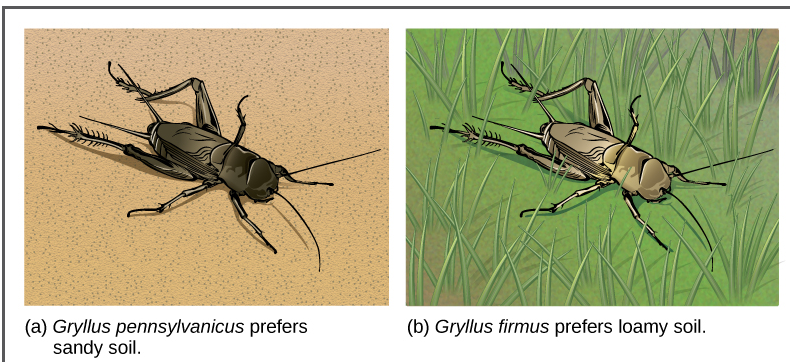
(a)

(b)

These two related frog species exhibit temporal reproductive isolation. (a) *Rana aurora* breeds earlier in the year than (b) *Rana boylei*. (credit a: modification of work by Mark R. Jennings, USFWS; credit b: modification of work by Alessandro Catenazzi)

In some cases, populations of a species move or are moved to a new habitat and take up residence in a place that no longer overlaps with the other

populations of the same species. This situation is called **habitat isolation**. Reproduction with the parent species ceases, and a new group exists that is now reproductively and genetically independent. For example, a cricket population that was divided after a flood could no longer interact with each other. Over time, the forces of natural selection, mutation, and genetic drift will likely result in the divergence of the two groups ([link](#)).



Speciation can occur when two populations occupy different habitats. The habitats need not be far apart. The cricket (a) *Gryllus pennsylvanicus* prefers sandy soil, and the cricket (b) *Gryllus firmus* prefers loamy soil. The two species can live in close proximity, but because of their different soil preferences, they became genetically isolated.

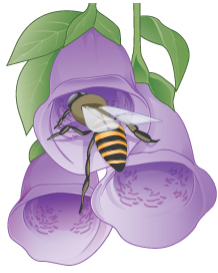
**Behavioral isolation** occurs when the presence or absence of a specific behavior prevents reproduction from taking place. For example, male fireflies use specific light patterns to attract females. Various species of fireflies display their lights differently. If a male of one species tried to attract the female of another, she would not recognize the light pattern and would not mate with the male.

Other prezygotic barriers work when differences in their gamete cells (eggs and sperm) prevent fertilization from taking place; this is called a **gametic barrier**. Similarly, in some cases closely related organisms try to mate, but their reproductive structures simply do not fit together. For example, damselfly males of different species have differently shaped reproductive organs. If one species tries to mate with the female of another, their body parts simply do not fit together. ([link](#)).



The shape of the male reproductive organ varies among male damselfly species, and is only compatible with the female of that species. Reproductive organ incompatibility keeps the species reproductively isolated.

In plants, certain structures aimed to attract one type of pollinator simultaneously prevent a different pollinator from accessing the pollen. The tunnel through which an animal must access nectar can vary widely in length and diameter, which prevents the plant from being cross-pollinated with a different species ([link](#)).



(a) Honeybee drinking nectar from a foxglove flower



(b) Ruby-throated hummingbird drinking nectar from a trumpet creeper flower

Some flowers have evolved to attract certain pollinators. The (a) wide foxglove flower is adapted for pollination by bees, while the (b) long, tube-shaped trumpet creeper flower is adapted for pollination by humming birds.

When fertilization takes place and a zygote forms, postzygotic barriers can prevent reproduction. Hybrid individuals in many cases cannot form normally in the womb and simply do not survive past the embryonic stages. This is called **hybrid inviability** because the hybrid organisms simply are not viable. In another postzygotic situation, reproduction leads to the birth and growth of a hybrid that is sterile and unable to reproduce offspring of their own; this is called hybrid sterility.

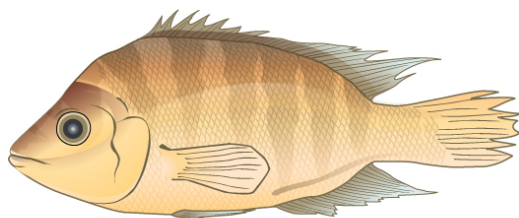
## Habitat Influence on Speciation

Sympatric speciation may also take place in ways other than polyploidy. For example, consider a species of fish that lives in a lake. As the population grows, competition for food also grows. Under pressure to find food, suppose that a group of these fish had the genetic flexibility to discover and feed off another resource that was unused by the other fish. What if this new food source was found at a different depth of the lake? Over time, those feeding on the second food source would interact more with each other than the other fish; therefore, they would breed together as well. Offspring of these fish would likely behave as their parents: feeding

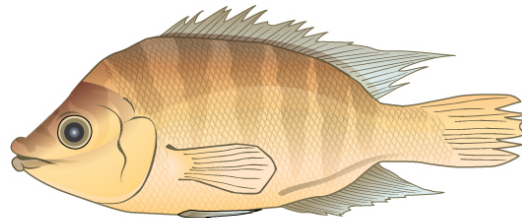


and living in the same area and keeping separate from the original population. If this group of fish continued to remain separate from the first population, eventually sympatric speciation might occur as more genetic differences accumulated between them.

This scenario does play out in nature, as do others that lead to reproductive isolation. One such place is Lake Victoria in Africa, famous for its sympatric speciation of cichlid fish. Researchers have found hundreds of sympatric speciation events in these fish, which have not only happened in great number, but also over a short period of time. [\[link\]](#) shows this type of speciation among a cichlid fish population in Nicaragua. In this locale, two types of cichlids live in the same geographic location but have come to have different morphologies that allow them to eat various food sources.



Thin-lipped cichlid



Thick-lipped cichlid

Cichlid fish from Lake Apoyeque, Nicaragua, show evidence of sympatric speciation. Lake Apoyeque, a crater lake, is 1800 years old, but genetic evidence indicates that the lake was populated only 100 years ago by a single population of cichlid fish. Nevertheless, two populations with distinct morphologies and diets now exist in the lake, and scientists believe these populations may be in an early stage of speciation.

## Section Summary

Speciation occurs along two main pathways: geographic separation (allopatric speciation) and through mechanisms that occur within a shared habitat (sympatric speciation). Both pathways isolate a population

reproductively in some form. Mechanisms of reproductive isolation act as barriers between closely related species, enabling them to diverge and exist as genetically independent species. Prezygotic barriers block reproduction prior to formation of a zygote, whereas postzygotic barriers block reproduction after fertilization occurs. For a new species to develop, something must cause a breach in the reproductive barriers. Sympatric speciation can occur through errors in meiosis that form gametes with extra chromosomes (polyploidy). Autopolyploidy occurs within a single species, whereas allopolyploidy occurs between closely related species.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Which is most likely to survive, offspring with  $2n+1$  chromosomes or offspring with  $2n-1$  chromosomes?

---

#### Solution:

[\[link\]](#) Loss of genetic material is almost always lethal, so offspring with  $2n+1$  chromosomes are more likely to survive.

## Review Questions

### Exercise:

#### Problem:

Which situation would most likely lead to allopatric speciation?

- a. flood causes the formation of a new lake.
- b. A storm causes several large trees to fall down.
- c. A mutation causes a new trait to develop.
- d. An injury causes an organism to seek out a new food source.



---

**Solution:**

A

**Exercise:**

**Problem:**

What is the main difference between dispersal and vicariance?

- a. One leads to allopatric speciation, whereas the other leads to sympatric speciation.
- b. One involves the movement of the organism, and the other involves a change in the environment.
- c. One depends on a genetic mutation occurring, and the other does not.
- d. One involves closely related organisms, and the other involves only individuals of the same species.

---

**Solution:**

B

**Exercise:**

**Problem:**

Which variable increases the likelihood of allopatric speciation taking place more quickly?

- a. lower rate of mutation
- b. longer distance between divided groups
- c. increased instances of hybrid formation
- d. equivalent numbers of individuals in each population

---

**Solution:**

B

**Exercise:**

**Problem:**

What is the main difference between autopolyploid and allopolyploid?

- a. the number of chromosomes
- b. the functionality of the chromosomes
- c. the source of the extra chromosomes
- d. the number of mutations in the extra chromosomes

---

**Solution:**

C

**Exercise:**

**Problem:** Which reproductive combination produces hybrids?

- a. when individuals of the same species in different geographical areas reproduce
- b. when any two individuals sharing the same habitat reproduce
- c. when members of closely related species reproduce
- d. when offspring of the same parents reproduce

---

**Solution:**

C

**Exercise:**

**Problem:**

Which condition is the basis for a species to be reproductively isolated from other members?

- a. It does not share its habitat with related species.
- b. It does not exist out of a single habitat.

- c. It does not exchange genetic information with other species.
- d. It does not undergo evolutionary changes for a significant period of time.

---

**Solution:**

C

**Exercise:**

**Problem:** Which situation is *not* an example of a prezygotic barrier?

- a. Two species of turtles breed at different times of the year.
- b. Two species of flowers attract different pollinators.
- c. Two species of birds display different mating dances.
- d. Two species of insects produce infertile offspring.

---

**Solution:**

D

## Free Response

**Exercise:**

**Problem:**

Why do island chains provide ideal conditions for adaptive radiation to occur?

---

**Solution:**

Organisms of one species can arrive to an island together and then disperse throughout the chain, each settling into different niches and exploiting different food resources to reduce competition.

**Exercise:**

**Problem:**

Two species of fish had recently undergone sympatric speciation. The males of each species had a different coloring through which the females could identify and choose a partner from her own species. After some time, pollution made the lake so cloudy that it was hard for females to distinguish colors. What might take place in this situation?

---

**Solution:**

It is likely the two species would start to reproduce with each other. Depending on the viability of their offspring, they may fuse back into one species.

**Exercise:****Problem:**

Why can polyploidy individuals lead to speciation fairly quickly?

---

**Solution:**

The formation of gametes with new  $n$  numbers can occur in one generation. After a couple of generations, enough of these new hybrids can form to reproduce together as a new species.

**Glossary**

adaptive radiation

speciation when one species radiates out to form several other species

allopatric speciation

speciation that occurs via geographic separation

allopolyploid

polyploidy formed between two related, but separate species

aneuploidy

condition of a cell having an extra chromosome or missing a chromosome for its species

autopolyploid

polyploidy formed within a single species

behavioral isolation

type of reproductive isolation that occurs when a specific behavior or lack of one prevents reproduction from taking place

dispersal

allopatric speciation that occurs when a few members of a species move to a new geographical area

gametic barrier

prezygotic barrier occurring when closely related individuals of different species mate, but differences in their gamete cells (eggs and sperm) prevent fertilization from taking place

habitat isolation

reproductive isolation resulting when populations of a species move or are moved to a new habitat, taking up residence in a place that no longer overlaps with the other populations of the same species

hybrid

offspring of two closely related individuals, not of the same species

postzygotic barrier

reproductive isolation mechanism that occurs after zygote formation

prezygotic barrier

reproductive isolation mechanism that occurs before zygote formation

reproductive isolation

situation that occurs when a species is reproductively independent from other species; this may be brought about by behavior, location, or reproductive barriers

speciation

formation of a new species

species

group of populations that interbreed and produce fertile offspring

sympatric speciation

speciation that occurs in the same geographic space

temporal isolation

differences in breeding schedules that can act as a form of prezygotic barrier leading to reproductive isolation

vicariance

allopatric speciation that occurs when something in the environment separates organisms of the same species into separate groups

## Reconnection and Rates of Speciation

By the end of this section, you will be able to:

- Describe pathways of species evolution in hybrid zones
- Explain the two major theories on rates of speciation

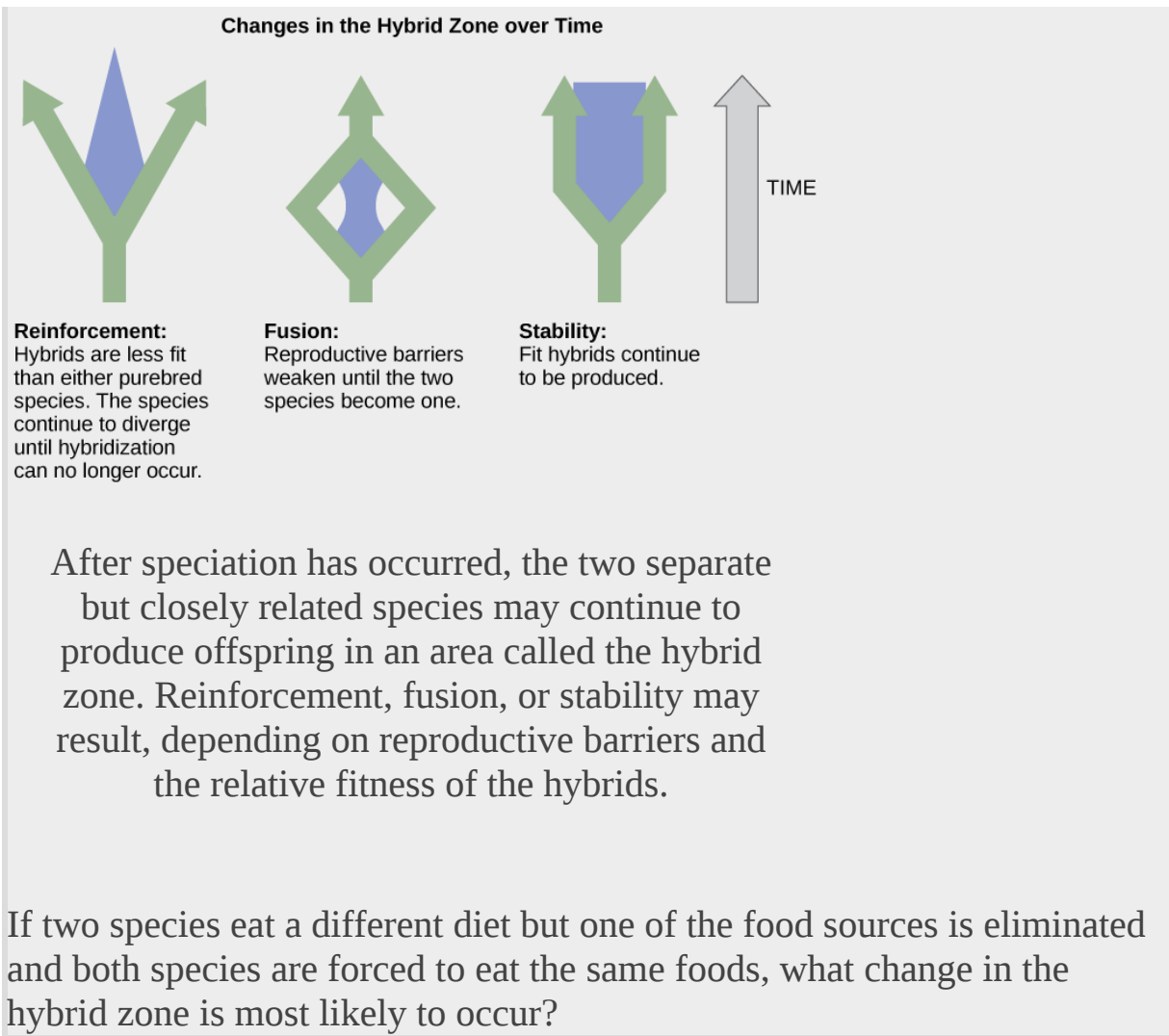
Speciation occurs over a span of evolutionary time, so when a new species arises, there is a transition period during which the closely related species continue to interact.

### Reconnection

After speciation, two species may recombine or even continue interacting indefinitely. Individual organisms will mate with any nearby individual who they are capable of breeding with. An area where two closely related species continue to interact and reproduce, forming hybrids, is called a **hybrid zone**. Over time, the hybrid zone may change depending on the fitness of the hybrids and the reproductive barriers ([link](#)). If the hybrids are less fit than the parents, reinforcement of speciation occurs, and the species continue to diverge until they can no longer mate and produce viable offspring. If reproductive barriers weaken, fusion occurs and the two species become one. Barriers remain the same if hybrids are fit and reproductive: stability may occur and hybridization continues.

#### **Note:**

Art Connection



Hybrids can be either less fit than the parents, more fit, or about the same. Usually hybrids tend to be less fit; therefore, such reproduction diminishes over time, nudging the two species to diverge further in a process called **reinforcement**. This term is used because the low success of the hybrids reinforces the original speciation. If the hybrids are as fit or more fit than the parents, the two species may fuse back into one species ([link](#)). Scientists have also observed that sometimes two species will remain separate but also continue to interact to produce some hybrid individuals; this is classified as stability because no real net change is taking place.

## Varying Rates of Speciation



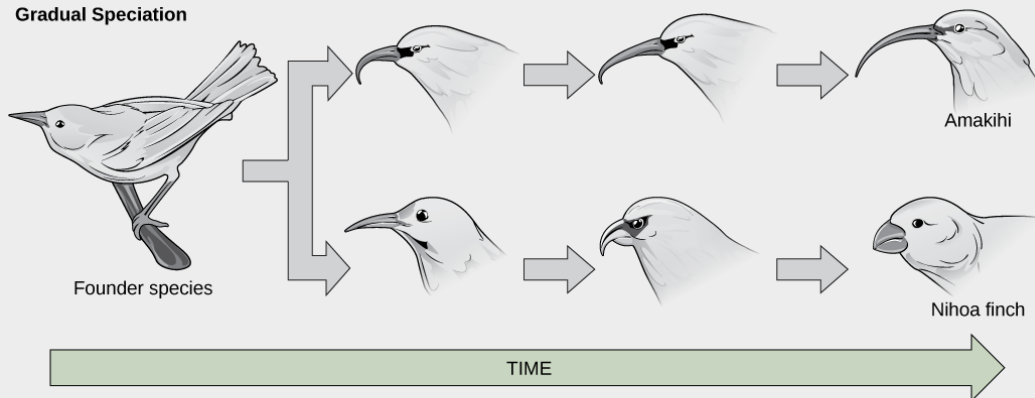
Scientists around the world study speciation, documenting observations both of living organisms and those found in the fossil record. As their ideas take shape and as research reveals new details about how life evolves, they develop models to help explain rates of speciation. In terms of how quickly speciation occurs, two patterns are currently observed: gradual speciation model and punctuated equilibrium model.

In the **gradual speciation model**, species diverge gradually over time in small steps. In the **punctuated equilibrium** model, a new species undergoes changes quickly from the parent species, and then remains largely unchanged for long periods of time afterward ([\[link\]](#)). This early change model is called punctuated equilibrium, because it begins with a punctuated or periodic change and then remains in balance afterward. While punctuated equilibrium suggests a faster tempo, it does not necessarily exclude gradualism.

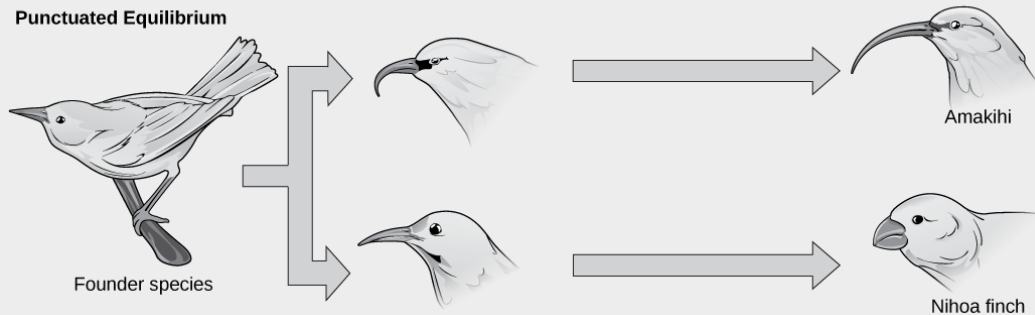
### Note:

### Art Connection

#### Gradual Speciation



#### Punctuated Equilibrium



In (a) gradual speciation, species diverge at a slow, steady pace as traits change incrementally. In (b) punctuated equilibrium, species diverge quickly and then remain unchanged for long periods of time.

Which of the following statements is false?

- a. Punctuated equilibrium is most likely to occur in a small population that experiences a rapid change in its environment.
- b. Punctuated equilibrium is most likely to occur in a large population that lives in a stable climate.
- c. Gradual speciation is most likely to occur in species that live in a stable climate.
- d. Gradual speciation and punctuated equilibrium both result in the divergence of species.

The primary influencing factor on changes in speciation rate is environmental conditions. Under some conditions, selection occurs quickly or radically. Consider a species of snails that had been living with the same basic form for many thousands of years. Layers of their fossils would appear similar for a long time. When a change in the environment takes place—such as a drop in the water level—a small number of organisms are separated from the rest in a brief period of time, essentially forming one large and one tiny population. The tiny population faces new environmental conditions. Because its gene pool quickly became so small, any variation that surfaces and that aids in surviving the new conditions becomes the predominant form.

**Note:**

[Link to Learning](#)



Visit [this website](#) to continue the speciation story of the snails.

## Section Summary

Speciation is not a precise division: overlap between closely related species can occur in areas called hybrid zones. Organisms reproduce with other similar organisms. The fitness of these hybrid offspring can affect the evolutionary path of the two species. Scientists propose two models for the rate of speciation: one model illustrates how a species can change slowly over time; the other model demonstrates how change can occur quickly from a parent generation to a new species. Both models continue to follow the patterns of natural selection.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) If two species eat a different diet but one of the food sources is eliminated and both species are forced to eat the same foods, what change in the hybrid zone is most likely to occur?

---

#### Solution:

[\[link\]](#) Fusion is most likely to occur because the two species will interact more and similar traits in food acquisition will be selected.

### Exercise:

**Problem:** [\[link\]](#) Which of the following statements is false?

- a. Punctuated equilibrium is most likely to occur in a small population that experiences a rapid change in its environment.
- b. Punctuated equilibrium is most likely to occur in a large population that lives in a stable climate.
- c. Gradual speciation is most likely to occur in species that live in a stable climate.
- d. Gradual speciation and punctuated equilibrium both result in the evolution of new species.

---

**Solution:**

[\[link\]](#) Answer B

## Review Questions

**Exercise:**

**Problem:**

Which term is used to describe the continued divergence of species based on the low fitness of hybrid offspring?

- a. reinforcement
- b. fusion
- c. stability
- d. punctuated equilibrium

---

**Solution:**

A

**Exercise:**

**Problem:**

Which components of speciation would be least likely to be a part of punctuated equilibrium?

- a. a division of populations
- b. a change in environmental conditions
- c. ongoing gene flow among all individuals
- d. a large number of mutations taking place at once

---

**Solution:**

C

**Free Response****Exercise:**

**Problem:** What do both rate of speciation models have in common?

---

**Solution:**

Both models continue to conform to the rules of natural selection, and the influences of gene flow, genetic drift, and mutation.

**Exercise:****Problem:**

Describe a situation where hybrid reproduction would cause two species to fuse into one.

---

**Solution:**

If the hybrid offspring are as fit or more fit than the parents, reproduction would likely continue between both species and the

hybrids, eventually bringing all organisms under the umbrella of one species.

## **Glossary**

gradual speciation model

model that shows how species diverge gradually over time in small steps

hybrid zone

area where two closely related species continue to interact and reproduce, forming hybrids

punctuated equilibrium

model for rapid speciation that can occur when an event causes a small portion of a population to be cut off from the rest of the population

reinforcement

continued speciation divergence between two related species due to low fitness of hybrids between them

## Introduction

class="introduction"

Living things may be single-celled or complex, multicellular organisms. They may be plants, animals, fungi, bacteria, or archaea. This diversity results from evolution.

(credit "wolf": modification of work by Gary Kramer; credit "coral": modification of work by William Harrigan, NOAA; credit "river": modification of work by Vojtěch

Dostál;  
credit "fish"  
modification  
of work by  
Christian  
Mehlführer;  
credit  
"mushroom"  
:  
modification  
of work by  
Cory  
Zanker;  
credit "tree":  
modification  
of work by  
Joseph  
Kranak;  
credit "bee":  
modification  
of work by  
Cory  
Zanker)





All life on Earth is related. Evolutionary theory states that humans, beetles, plants, and bacteria all share a common ancestor, but that millions of years of evolution have shaped each of these organisms into the forms seen today. Scientists consider evolution a key concept to understanding life. Natural selection is one of the most dominant evolutionary forces. Natural selection acts to promote traits and behaviors that increase an organism's chances of survival and reproduction, while eliminating those traits and behaviors that are to the organism's detriment. But natural selection can only, as its name implies, select—it cannot create. The introduction of novel traits and behaviors falls on the shoulders of another evolutionary force—mutation. Mutation and other sources of variation among individuals, as well as the evolutionary forces that act upon them, alter populations and species. This combination of processes has led to the world of life we see today.

## Population Evolution

By the end of this section, you will be able to:

- Define population genetics and describe how population genetics is used in the study of the evolution of populations
- Define the Hardy-Weinberg principle and discuss its importance

The mechanisms of inheritance, or genetics, were not understood at the time Charles Darwin and Alfred Russel Wallace were developing their idea of natural selection. This lack of understanding was a stumbling block to understanding many aspects of evolution. In fact, the predominant (and incorrect) genetic theory of the time, blending inheritance, made it difficult to understand how natural selection might operate. Darwin and Wallace were unaware of the genetics work by Austrian monk Gregor Mendel, which was published in 1866, not long after publication of Darwin's book, *On the Origin of Species*. Mendel's work was rediscovered in the early twentieth century at which time geneticists were rapidly coming to an understanding of the basics of inheritance. Initially, the newly discovered particulate nature of genes made it difficult for biologists to understand how gradual evolution could occur. But over the next few decades genetics and evolution were integrated in what became known as the **modern synthesis**—the coherent understanding of the relationship between natural selection and genetics that took shape by the 1940s and is generally accepted today. In sum, the modern synthesis describes how evolutionary processes, such as natural selection, can affect a population's genetic makeup, and, in turn, how this can result in the gradual evolution of populations and species. The theory also connects this change of a population over time, called **microevolution**, with the processes that gave rise to new species and higher taxonomic groups with widely divergent characters, called **macroevolution**.

### Note:

#### Everyday Connection

#### Evolution and Flu Vaccines

Every fall, the media starts reporting on flu vaccinations and potential outbreaks. Scientists, health experts, and institutions determine

recommendations for different parts of the population, predict optimal production and inoculation schedules, create vaccines, and set up clinics to provide inoculations. You may think of the annual flu shot as a lot of media hype, an important health protection, or just a briefly uncomfortable prick in your arm. But do you think of it in terms of evolution?

The media hype of annual flu shots is scientifically grounded in our understanding of evolution. Each year, scientists across the globe strive to predict the flu strains that they anticipate being most widespread and harmful in the coming year. This knowledge is based in how flu strains have evolved over time and over the past few flu seasons. Scientists then work to create the most effective vaccine to combat those selected strains. Hundreds of millions of doses are produced in a short period in order to provide vaccinations to key populations at the optimal time.

Because viruses, like the flu, evolve very quickly (especially in evolutionary time), this poses quite a challenge. Viruses mutate and replicate at a fast rate, so the vaccine developed to protect against last year's flu strain may not provide the protection needed against the coming year's strain. Evolution of these viruses means continued adaptations to ensure survival, including adaptations to survive previous vaccines.

## Population Genetics

Recall that a gene for a particular character may have several alleles, or variants, that code for different traits associated with that character. For example, in the ABO blood type system in humans, three alleles determine the particular blood-type carbohydrate on the surface of red blood cells. Each individual in a population of diploid organisms can only carry two alleles for a particular gene, but more than two may be present in the individuals that make up the population. Mendel followed alleles as they were inherited from parent to offspring. In the early twentieth century, biologists in a field of study known as **population genetics** began to study how selective forces change a population through changes in allele and genotypic frequencies.

The **allele frequency** (or gene frequency) is the rate at which a specific allele appears within a population. Until now we have discussed evolution as a change in the characteristics of a population of organisms, but behind that phenotypic change is genetic change. In population genetics, the term evolution is defined as a change in the frequency of an allele in a population. Using the ABO blood type system as an example, the frequency of one of the alleles,  $I^A$ , is the number of copies of that allele divided by all the copies of the ABO gene in the population. For example, a study in Jordan<sup>[footnote]</sup> found a frequency of  $I^A$  to be 26.1 percent. The  $I^B$  and  $I^O$  alleles made up 13.4 percent and 60.5 percent of the alleles respectively, and all of the frequencies added up to 100 percent. A change in this frequency over time would constitute evolution in the population. Sahar S. Hanania, Dhia S. Hassawi, and Nidal M. Irshaid, “Allele Frequency and Molecular Genotypes of ABO Blood Group System in a Jordanian Population,” *Journal of Medical Sciences* 7 (2007): 51-58, doi:10.3923/jms.2007.51.58.

The allele frequency within a given population can change depending on environmental factors; therefore, certain alleles become more widespread than others during the process of natural selection. Natural selection can alter the population’s genetic makeup; for example, if a given allele confers a phenotype that allows an individual to better survive or have more offspring. Because many of those offspring will also carry the beneficial allele, and often the corresponding phenotype, they will have more offspring of their own that also carry the allele, thus, perpetuating the cycle. Over time, the allele will spread throughout the population. Some alleles will quickly become fixed in this way, meaning that every individual of the population will carry the allele, while detrimental mutations may be swiftly eliminated if derived from a dominant allele from the gene pool. The **gene pool** is the sum of all the alleles in a population.

Sometimes, allele frequencies within a population change randomly with no advantage to the population over existing allele frequencies. This phenomenon is called genetic drift. Natural selection and genetic drift usually occur simultaneously in populations and are not isolated events. It is hard to determine which process dominates because it is often nearly impossible to determine the cause of change in allele frequencies at each

occurrence. An event that initiates an allele frequency change in an isolated part of the population, which is not typical of the original population, is called the **founder effect**. Natural selection, random drift, and founder effects can lead to significant changes in the genome of a population.

## Hardy-Weinberg Principle of Equilibrium

In the early twentieth century, English mathematician Godfrey Hardy and German physician Wilhelm Weinberg stated the principle of equilibrium to describe the genetic makeup of a population. The theory, which later became known as the Hardy-Weinberg principle of equilibrium, states that a population's allele and genotype frequencies are inherently stable— unless some kind of evolutionary force is acting upon the population, neither the allele nor the genotypic frequencies would change. The Hardy-Weinberg principle assumes conditions with no mutations, migration, emigration, or selective pressure for or against genotype, plus an infinite population; while no population can satisfy those conditions, the principle offers a useful model against which to compare real population changes.

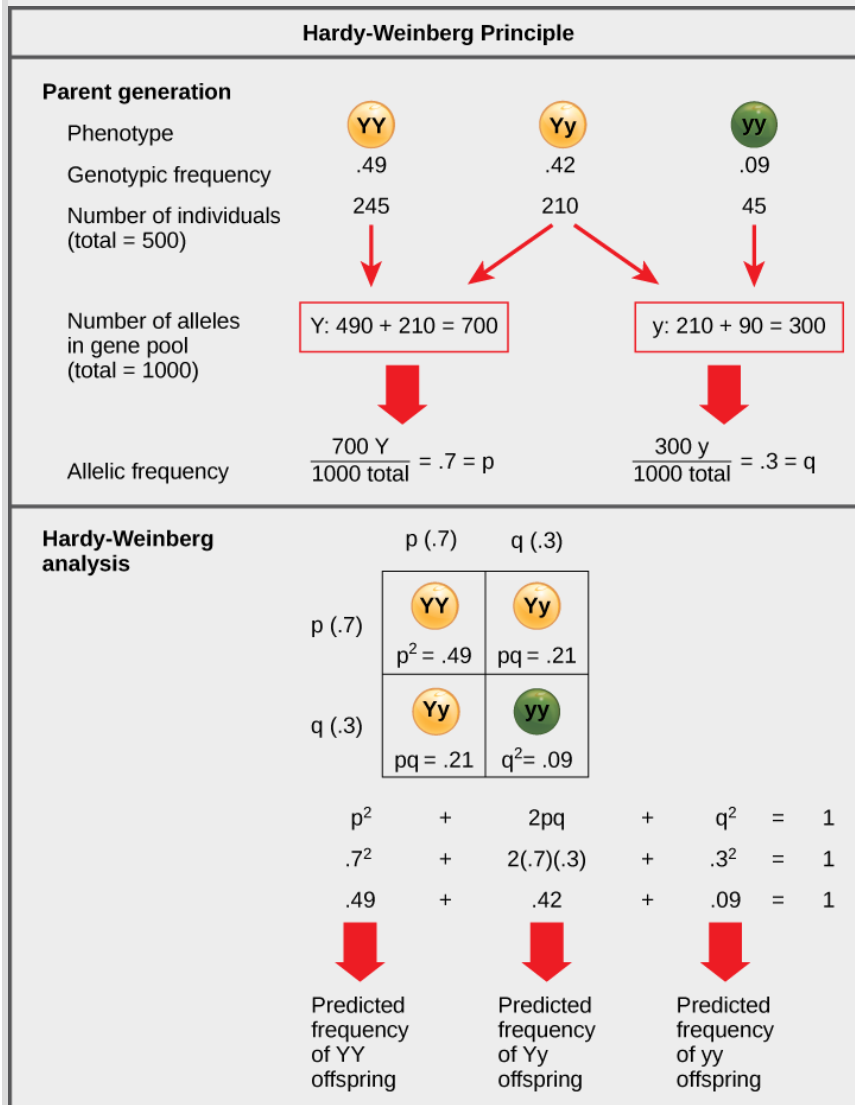
Working under this theory, population geneticists represent different alleles as different variables in their mathematical models. The variable  $p$ , for example, often represents the frequency of a particular allele, say  $Y$  for the trait of yellow in Mendel's peas, while the variable  $q$  represents the frequency of  $y$  alleles that confer the color green. If these are the only two possible alleles for a given locus in the population,  $p + q = 1$ . In other words, all the  $p$  alleles and all the  $q$  alleles make up all of the alleles for that locus that are found in the population.

But what ultimately interests most biologists is not the frequencies of different alleles, but the frequencies of the resulting genotypes, known as the population's **genetic structure**, from which scientists can surmise the distribution of phenotypes. If the phenotype is observed, only the genotype of the homozygous recessive alleles can be known; the calculations provide an estimate of the remaining genotypes. Since each individual carries two alleles per gene, if the allele frequencies ( $p$  and  $q$ ) are known, predicting the frequencies of these genotypes is a simple mathematical calculation to determine the probability of getting these genotypes if two alleles are drawn

at random from the gene pool. So in the above scenario, an individual pea plant could be pp (YY), and thus produce yellow peas; pq (Yy), also yellow; or qq (yy), and thus producing green peas ([\[link\]](#)). In other words, the frequency of pp individuals is simply  $p^2$ ; the frequency of pq individuals is  $2pq$ ; and the frequency of qq individuals is  $q^2$ . And, again, if p and q are the only two possible alleles for a given trait in the population, these genotypes frequencies will sum to one:  $p^2 + 2pq + q^2 = 1$ .

### Note:

### Art Connection



When populations are in the Hardy-Weinberg equilibrium, the allelic frequency is stable from generation to generation and the distribution of alleles can be determined from the Hardy-Weinberg equation. If the allelic frequency measured in the field differs from the predicted value, scientists can make inferences about what evolutionary forces are at play.

In plants, violet flower color (V) is dominant over white (v). If  $p = 0.8$  and  $q = 0.2$  in a population of 500 plants, how many individuals would you expect to be homozygous dominant (VV), heterozygous (Vv), and homozygous recessive (vv)? How many plants would you expect to have violet flowers, and how many would have white flowers?

In theory, if a population is at equilibrium—that is, there are no evolutionary forces acting upon it—generation after generation would have the same gene pool and genetic structure, and these equations would all hold true all of the time. Of course, even Hardy and Weinberg recognized that no natural population is immune to evolution. Populations in nature are constantly changing in genetic makeup due to drift, mutation, possibly migration, and selection. As a result, the only way to determine the exact distribution of phenotypes in a population is to go out and count them. But the Hardy-Weinberg principle gives scientists a mathematical baseline of a non-evolving population to which they can compare evolving populations and thereby infer what evolutionary forces might be at play. If the frequencies of alleles or genotypes deviate from the value expected from the Hardy-Weinberg equation, then the population is evolving.

**Note:**

[Link to Learning](#)



Use this [online calculator](#) to determine the genetic structure of a population.

## Section Summary

The modern synthesis of evolutionary theory grew out of the cohesion of Darwin's, Wallace's, and Mendel's thoughts on evolution and heredity, along with the more modern study of population genetics. It describes the evolution of populations and species, from small-scale changes among individuals to large-scale changes over paleontological time periods. To understand how organisms evolve, scientists can track populations' allele frequencies over time. If they differ from generation to generation, scientists can conclude that the population is not in Hardy-Weinberg equilibrium, and is thus evolving.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) In plants, violet flower color (V) is dominant over white (v). If  $p = .8$  and  $q = 0.2$  in a population of 500 plants, how many individuals would you expect to be homozygous dominant (VV), heterozygous (Vv), and homozygous recessive (vv)? How many plants would you expect to have violet flowers, and how many would have white flowers?

---

#### Solution:



[\[link\]](#) The expected distribution is 320 VV, 160Vv, and 20 vv plants. Plants with VV or Vv genotypes would have violet flowers, and plants with the vv genotype would have white flowers, so a total of 480 plants would be expected to have violet flowers, and 20 plants would have white flowers.

## Review Questions

### Exercise:

**Problem:** What is the difference between micro- and macroevolution?

- a. Microevolution describes the evolution of small organisms, such as insects, while macroevolution describes the evolution of large organisms, like people and elephants.
- b. Microevolution describes the evolution of microscopic entities, such as molecules and proteins, while macroevolution describes the evolution of whole organisms.
- c. Microevolution describes the evolution of organisms in populations, while macroevolution describes the evolution of species over long periods of time.
- d. Microevolution describes the evolution of organisms over their lifetimes, while macroevolution describes the evolution of organisms over multiple generations.

---

### Solution:

C

### Exercise:

**Problem:** Population genetics is the study of:

- a. how selective forces change the allele frequencies in a population over time
- b. the genetic basis of population-wide traits

- c. whether traits have a genetic basis
  - d. the degree of inbreeding in a population
- 

**Solution:**

A

**Exercise:**

**Problem:**

Which of the following populations is not in Hardy-Weinberg equilibrium?

- a. a population with 12 homozygous recessive individuals (yy), 8 homozygous dominant individuals (YY), and 4 heterozygous individuals (Yy)
  - b. a population in which the allele frequencies do not change over time
  - c.  $p^2 + 2pq + q^2 = 1$
  - d. a population undergoing natural selection
- 

**Solution:**

D

**Exercise:**

**Problem:**

One of the original Amish colonies rose from a ship of colonists that came from Europe. The ship's captain, who had polydactyly, a rare dominant trait, was one of the original colonists. Today, we see a much higher frequency of polydactyly in the Amish population. This is an example of:

- a. natural selection
- b. genetic drift

- c. founder effect
- d. b and c

---

**Solution:**

D

**Free Response**

**Exercise:**

**Problem:**

Solve for the genetic structure of a population with 12 homozygous recessive individuals (yy), 8 homozygous dominant individuals (YY), and 4 heterozygous individuals (Yy).

---

**Solution:**

$$p = (8 \cdot 2 + 4) / 48 = .42; q = (12 \cdot 2 + 4) / 48 = .58; p^2 = .17; 2pq = .48; q^2 = .34$$

**Exercise:**

**Problem:** Explain the Hardy-Weinberg principle of equilibrium theory.

---

**Solution:**

The Hardy-Weinberg principle of equilibrium is used to describe the genetic makeup of a population. The theory states that a population's allele and genotype frequencies are inherently stable: unless some kind of evolutionary force is acting upon the population, generation after generation of the population would carry the same genes, and individuals would, as a whole, look essentially the same.

**Exercise:**

**Problem:**

Imagine you are trying to test whether a population of flowers is undergoing evolution. You suspect there is selection pressure on the color of the flower: bees seem to cluster around the red flowers more often than the blue flowers. In a separate experiment, you discover blue flower color is dominant to red flower color. In a field, you count 600 blue flowers and 200 red flowers. What would you expect the genetic structure of the flowers to be?

---

**Solution:**

Red is recessive so  $q^2 = 200/800 = 0.25$ ;  $q = 0.5$ ;  $p = 1 - q = 0.5$ ;  $p^2 = 0.25$ ;  $2pq = 0.5$ . You would expect 200 homozygous blue flowers, 400 heterozygous blue flowers, and 200 red flowers.

**Glossary**

allele frequency

(also, gene frequency) rate at which a specific allele appears within a population

founder effect

event that initiates an allele frequency change in part of the population, which is not typical of the original population

gene pool

all of the alleles carried by all of the individuals in the population

genetic structure

distribution of the different possible genotypes in a population

macroevolution

broader scale evolutionary changes seen over paleontological time

microevolution

changes in a population's genetic structure

modern synthesis

overarching evolutionary paradigm that took shape by the 1940s and is generally accepted today

population genetics

study of how selective forces change the allele frequencies in a population over time

## Population Genetics

By the end of this section, you will be able to:

- Describe the different types of variation in a population
- Explain why only heritable variation can be acted upon by natural selection
- Describe genetic drift and the bottleneck effect
- Explain how each evolutionary force can influence the allele frequencies of a population

Individuals of a population often display different phenotypes, or express different alleles of a particular gene, referred to as polymorphisms.

Populations with two or more variations of particular characteristics are called polymorphic. The distribution of phenotypes among individuals, known as the **population variation**, is influenced by a number of factors, including the population's genetic structure and the environment ([\[link\]](#)).

Understanding the sources of a phenotypic variation in a population is important for determining how a population will evolve in response to different evolutionary pressures.



The distribution of phenotypes in this litter of kittens illustrates population variation. (credit: Pieter Lanser)

## Genetic Variance

Natural selection and some of the other evolutionary forces can only act on heritable traits, namely an organism's genetic code. Because alleles are passed from parent to offspring, those that confer beneficial traits or behaviors may be selected for, while deleterious alleles may be selected against. Acquired traits, for the most part, are not heritable. For example, if an athlete works out in the gym every day, building up muscle strength, the athlete's offspring will not necessarily grow up to be a body builder. If there is a genetic basis for the ability to run fast, on the other hand, this may be passed to a child.

**Note:**

Link to Learning



Before Darwinian evolution became the prevailing theory of the field, French naturalist Jean-Baptiste Lamarck theorized that acquired traits could, in fact, be inherited; while this hypothesis has largely been unsupported, scientists have recently begun to realize that Lamarck was not completely wrong. Visit this [site](#) to learn more.

**Heritability** is the fraction of phenotype variation that can be attributed to genetic differences, or genetic variance, among individuals in a population. The greater the heritability of a population's phenotypic variation, the more susceptible it is to the evolutionary forces that act on heritable variation.

The diversity of alleles and genotypes within a population is called **genetic variance**. When scientists are involved in the breeding of a species, such as with animals in zoos and nature preserves, they try to increase a

population's genetic variance to preserve as much of the phenotypic diversity as they can. This also helps reduce the risks associated with **inbreeding**, the mating of closely related individuals, which can have the undesirable effect of bringing together deleterious recessive mutations that can cause abnormalities and susceptibility to disease. For example, a disease that is caused by a rare, recessive allele might exist in a population, but it will only manifest itself when an individual carries two copies of the allele. Because the allele is rare in a normal, healthy population with unrestricted habitat, the chance that two carriers will mate is low, and even then, only 25 percent of their offspring will inherit the disease allele from both parents. While it is likely to happen at some point, it will not happen frequently enough for natural selection to be able to swiftly eliminate the allele from the population, and as a result, the allele will be maintained at low levels in the gene pool. However, if a family of carriers begins to interbreed with each other, this will dramatically increase the likelihood of two carriers mating and eventually producing diseased offspring, a phenomenon known as **inbreeding depression**.

Changes in allele frequencies that are identified in a population can shed light on how it is evolving. In addition to natural selection, there are other evolutionary forces that could be in play: genetic drift, gene flow, mutation, nonrandom mating, and environmental variances.

## **Genetic Drift**

The theory of natural selection stems from the observation that some individuals in a population are more likely to survive longer and have more offspring than others; thus, they will pass on more of their genes to the next generation. A big, powerful male gorilla, for example, is much more likely than a smaller, weaker one to become the population's silverback, the pack's leader who mates far more than the other males of the group. The pack leader will father more offspring, who share half of his genes, and are likely to also grow bigger and stronger like their father. Over time, the genes for bigger size will increase in frequency in the population, and the population will, as a result, grow larger on average. That is, this would occur if this particular **selection pressure**, or driving selective force, were



the only one acting on the population. In other examples, better camouflage or a stronger resistance to drought might pose a selection pressure.

Another way a population's allele and genotype frequencies can change is **genetic drift** ([\[link\]](#)), which is simply the effect of chance. By chance, some individuals will have more offspring than others—not due to an advantage conferred by some genetically-encoded trait, but just because one male happened to be in the right place at the right time (when the receptive female walked by) or because the other one happened to be in the wrong place at the wrong time (when a fox was hunting).

**Note:**

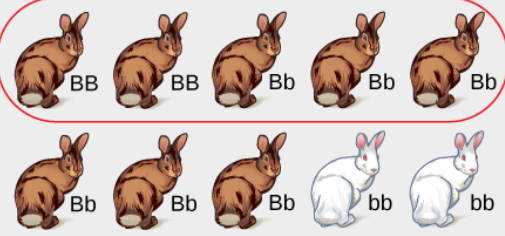
Art Connection

## Genetic Drift

### First generation

$p$  (B gene frequency) = .5

$q$  (b gene frequency) = .5



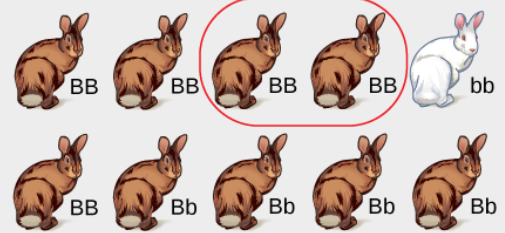
5 rabbits reproduce



### Second generation

$p$  = .7

$q$  = .3



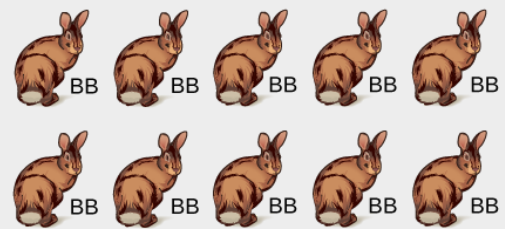
2 rabbits reproduce



### Third generation

$p$  = 1

$q$  = 0



Genetic drift in a population can lead to the elimination of an allele from a population by chance. In this example, rabbits with the brown coat color allele ( $B$ ) are dominant over rabbits with the white coat color allele ( $b$ ). In the

first generation, the two alleles occur with equal frequency in the population, resulting in  $p$  and  $q$  values of .5. Only half of the individuals reproduce, resulting in a second generation with  $p$  and  $q$  values of .7 and .3, respectively. Only two individuals in the second generation reproduce, and by chance these individuals are homozygous dominant for brown coat color. As a result, in the third generation the recessive  $b$  allele is lost.

Do you think genetic drift would happen more quickly on an island or on the mainland?

Small populations are more susceptible to the forces of genetic drift. Large populations, on the other hand, are buffered against the effects of chance. If one individual of a population of 10 individuals happens to die at a young age before it leaves any offspring to the next generation, all of its genes—1/10 of the population's gene pool—will be suddenly lost. In a population of 100, that's only 1 percent of the overall gene pool; therefore, it is much less impactful on the population's genetic structure.

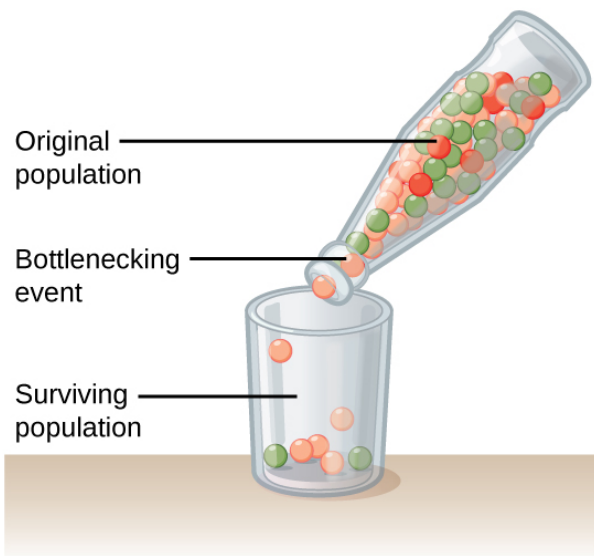
**Note:**

[Link to Learning](#)



Go to this [site](#) to watch an animation of random sampling and genetic drift in action.

Genetic drift can also be magnified by natural events, such as a natural disaster that kills—at random—a large portion of the population. Known as the **bottleneck effect**, it results in a large portion of the genome suddenly being wiped out ([link](#)). In one fell swoop, the genetic structure of the survivors becomes the genetic structure of the entire population, which may be very different from the pre-disaster population.



A chance event or catastrophe  
can reduce the genetic  
variability within a population.

Another scenario in which populations might experience a strong influence of genetic drift is if some portion of the population leaves to start a new population in a new location or if a population gets divided by a physical barrier of some kind. In this situation, those individuals are unlikely to be representative of the entire population, which results in the founder effect. The founder effect occurs when the genetic structure changes to match that of the new population's founding fathers and mothers. The founder effect is believed to have been a key factor in the genetic history of the Afrikaner population of Dutch settlers in South Africa, as evidenced by mutations that are common in Afrikaners but rare in most other populations. This is likely due to the fact that a higher-than-normal proportion of the founding colonists carried these mutations. As a result, the population expresses unusually high incidences of Huntington's disease (HD) and Fanconi anemia (FA), a genetic disorder known to cause blood marrow and congenital abnormalities—even cancer.<sup>[footnote]</sup>

A. J. Tipping et al., “Molecular and Genealogical Evidence for a Founder Effect in Fanconi Anemia Families of the Afrikaner Population of South Africa,” *PNAS* 98, no. 10 (2001): 5734-5739, doi: 10.1073/pnas.091402398.

**Note:**

Link to Learning



Watch this short video to learn more about the founder and bottleneck effects.

[https://www.openstaxcollege.org/l/founder\\_bottle](https://www.openstaxcollege.org/l/founder_bottle)

**Note:****Scientific Method Connection****Testing the Bottleneck Effect**

**Question:** How do natural disasters affect the genetic structure of a population?

**Background:** When much of a population is suddenly wiped out by an earthquake or hurricane, the individuals that survive the event are usually a random sampling of the original group. As a result, the genetic makeup of the population can change dramatically. This phenomenon is known as the bottleneck effect.

**Hypothesis:** Repeated natural disasters will yield different population genetic structures; therefore, each time this experiment is run, the results will vary.

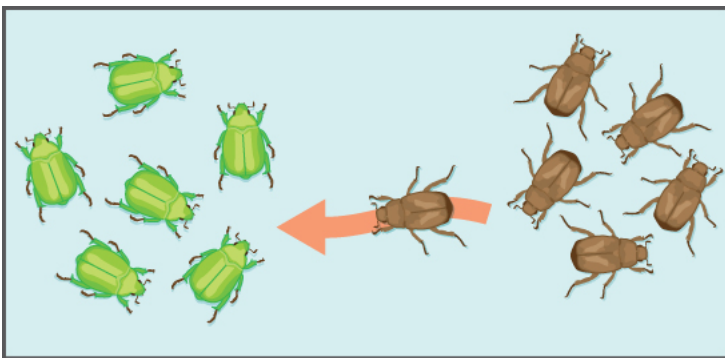
**Test the hypothesis:** Count out the original population using different colored beads. For example, red, blue, and yellow beads might represent red, blue, and yellow individuals. After recording the number of each individual in the original population, place them all in a bottle with a narrow neck that will only allow a few beads out at a time. Then, pour 1/3 of the bottle's contents into a bowl. This represents the surviving individuals after a natural disaster kills a majority of the population. Count the number of the different colored beads in the bowl, and record it. Then, place all of the beads back in the bottle and repeat the experiment four more times.

**Analyze the data:** Compare the five populations that resulted from the experiment. Do the populations all contain the same number of different colored beads, or do they vary? Remember, these populations all came from the same exact parent population.

**Form a conclusion:** Most likely, the five resulting populations will differ quite dramatically. This is because natural disasters are not selective—they kill and spare individuals at random. Now think about how this might affect a real population. What happens when a hurricane hits the Mississippi Gulf Coast? How do the seabirds that live on the beach fare?

**Gene Flow**

Another important evolutionary force is **gene flow**: the flow of alleles in and out of a population due to the migration of individuals or gametes ([link](#)). While some populations are fairly stable, others experience more flux. Many plants, for example, send their pollen far and wide, by wind or by bird, to pollinate other populations of the same species some distance away. Even a population that may initially appear to be stable, such as a pride of lions, can experience its fair share of immigration and emigration as developing males leave their mothers to seek out a new pride with genetically unrelated females. This variable flow of individuals in and out of the group not only changes the gene structure of the population, but it can also introduce new genetic variation to populations in different geological locations and habitats.



Gene flow can occur when an individual travels from one geographic location to another.

## Mutation

Mutations are changes to an organism's DNA and are an important driver of diversity in populations. Species evolve because of the accumulation of mutations that occur over time. The appearance of new mutations is the most common way to introduce novel genotypic and phenotypic variance. Some mutations are unfavorable or harmful and are quickly eliminated

from the population by natural selection. Others are beneficial and will spread through the population. Whether or not a mutation is beneficial or harmful is determined by whether it helps an organism survive to sexual maturity and reproduce. Some mutations do not do anything and can linger, unaffected by natural selection, in the genome. Some can have a dramatic effect on a gene and the resulting phenotype.

## **Nonrandom Mating**

If individuals nonrandomly mate with their peers, the result can be a changing population. There are many reasons **nonrandom mating** occurs. One reason is simple mate choice; for example, female peahens may prefer peacocks with bigger, brighter tails. Traits that lead to more matings for an individual become selected for by natural selection. One common form of mate choice, called **assortative mating**, is an individual's preference to mate with partners who are phenotypically similar to themselves.

Another cause of nonrandom mating is physical location. This is especially true in large populations spread over large geographic distances where not all individuals will have equal access to one another. Some might be miles apart through woods or over rough terrain, while others might live immediately nearby.

## **Environmental Variance**

Genes are not the only players involved in determining population variation. Phenotypes are also influenced by other factors, such as the environment ([link](#)). A beachgoer is likely to have darker skin than a city dweller, for example, due to regular exposure to the sun, an environmental factor. Some major characteristics, such as sex, are determined by the environment for some species. For example, some turtles and other reptiles have temperature-dependent sex determination (TSD). TSD means that individuals develop into males if their eggs are incubated within a certain temperature range, or females at a different temperature range.





The sex of the American alligator (*Alligator mississippiensis*) is determined by the temperature at which the eggs are incubated. Eggs incubated at 30°C produce females, and eggs incubated at 33°C produce males. (credit: Steve Hillebrand, USFWS)

Geographic separation between populations can lead to differences in the phenotypic variation between those populations. Such **geographical variation** is seen between most populations and can be significant. One type of geographic variation, called a **cline**, can be seen as populations of a given species vary gradually across an ecological gradient. Species of warm-blooded animals, for example, tend to have larger bodies in the cooler climates closer to the earth's poles, allowing them to better conserve heat. This is considered a latitudinal cline. Alternatively, flowering plants tend to bloom at different times depending on where they are along the slope of a mountain, known as an altitudinal cline.

If there is gene flow between the populations, the individuals will likely show gradual differences in phenotype along the cline. Restricted gene flow, on the other hand, can lead to abrupt differences, even speciation.

## Section Summary

Both genetic and environmental factors can cause phenotypic variation in a population. Different alleles can confer different phenotypes, and different environments can also cause individuals to look or act differently. Only those differences encoded in an individual's genes, however, can be passed to its offspring and, thus, be a target of natural selection. Natural selection works by selecting for alleles that confer beneficial traits or behaviors, while selecting against those for deleterious qualities. Genetic drift stems from the chance occurrence that some individuals in the germ line have more offspring than others. When individuals leave or join the population, allele frequencies can change as a result of gene flow. Mutations to an individual's DNA may introduce new variation into a population. Allele frequencies can also be altered when individuals do not randomly mate with others in the group.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Do you think genetic drift would happen more quickly on an island or on the mainland?

---

#### Solution:

[\[link\]](#) Genetic drift is likely to occur more rapidly on an island where smaller populations are expected to occur.

## Review Questions

### Exercise:

**Problem:**

When male lions reach sexual maturity, they leave their group in search of a new pride. This can alter the allele frequencies of the population through which of the following mechanisms?

- a. natural selection
- b. genetic drift
- c. gene flow
- d. random mating

---

**Solution:**

C

**Exercise:****Problem:**

Which of the following evolutionary forces can introduce new genetic variation into a population?

- a. natural selection and genetic drift
- b. mutation and gene flow
- c. natural selection and nonrandom mating
- d. mutation and genetic drift

---

**Solution:**

B

**Exercise:**

**Problem:** What is assortative mating?

- a. when individuals mate with those who are similar to themselves

- b. when individuals mate with those who are dissimilar to themselves
  - c. when individuals mate with those who are the most fit in the population
  - d. when individuals mate with those who are least fit in the population
- 

**Solution:**

A

**Exercise:**

**Problem:**

When closely related individuals mate with each other, or inbreed, the offspring are often not as fit as the offspring of two unrelated individuals. Why?

- a. Close relatives are genetically incompatible.
  - b. The DNA of close relatives reacts negatively in the offspring.
  - c. Inbreeding can bring together rare, deleterious mutations that lead to harmful phenotypes.
  - d. Inbreeding causes normally silent alleles to be expressed.
- 

**Solution:**

C

**Exercise:**

**Problem:**What is a cline?

- a. the slope of a mountain where a population lives
- b. the degree to which a mutation helps an individual survive
- c. the number of individuals in the population
- d. gradual geographic variation across an ecological gradient

---

**Solution:**

D

**Free Response****Exercise:****Problem:**

Describe a situation in which a population would undergo the bottleneck effect and explain what impact that would have on the population's gene pool.

---

**Solution:**

A hurricane kills a large percentage of a population of sand-dwelling crustaceans—only a few individuals survive. The alleles carried by those surviving individuals would represent the entire population's gene pool. If those surviving individuals are not representative of the original population, the post-hurricane gene pool will differ from the original gene pool.

**Exercise:****Problem:**

Describe natural selection and give an example of natural selection at work in a population.

---

**Solution:**

The theory of natural selection stems from the observation that some individuals in a population survive longer and have more offspring than others: thus, more of their genes are passed to the next generation. For example, a big, powerful male gorilla is much more likely than a smaller, weaker one to become the population's silverback: the pack's leader who mates far more than the other males of the group.

Therefore, the pack leader will father more offspring who share half of his genes and are likely to grow bigger and stronger like their father. Over time, the genes for bigger size will increase in frequency in the population, and the average body size, as a result, grow larger on average.

### **Exercise:**

**Problem:** Explain what a cline is and provide examples.

---

### **Solution:**

A cline is a type of geographic variation that is seen in populations of a given species that vary gradually across an ecological gradient. For example, warm-blooded animals tend to have larger bodies in the cooler climates closer to the earth's poles, allowing them to better conserve heat. This is considered a latitudinal cline. Flowering plants tend to bloom at different times depending on where they are along the slope of a mountain. This is known as an altitudinal cline.

## **Glossary**

assortative mating

when individuals tend to mate with those who are phenotypically similar to themselves

bottleneck effect

magnification of genetic drift as a result of natural events or catastrophes

cline

gradual geographic variation across an ecological gradient

gene flow

flow of alleles in and out of a population due to the migration of individuals or gametes

genetic drift

effect of chance on a population's gene pool

genetic variance

diversity of alleles and genotypes in a population

geographical variation

differences in the phenotypic variation between populations that are separated geographically

heritability

fraction of population variation that can be attributed to its genetic variance

inbreeding

mating of closely related individuals

inbreeding depression

increase in abnormalities and disease in inbreeding populations

nonrandom mating

changes in a population's gene pool due to mate choice or other forces that cause individuals to mate with certain phenotypes more than others

population variation

distribution of phenotypes in a population

selective pressure

environmental factor that causes one phenotype to be better than another

## Adaptive Evolution

By the end of this section, you will be able to:

- Explain the different ways natural selection can shape populations
- Describe how these different forces can lead to different outcomes in terms of the population variation

Natural selection only acts on the population's heritable traits: selecting for beneficial alleles and thus increasing their frequency in the population, while selecting against deleterious alleles and thereby decreasing their frequency—a process known as **adaptive evolution**. Natural selection does not act on individual alleles, however, but on entire organisms. An individual may carry a very beneficial genotype with a resulting phenotype that, for example, increases the ability to reproduce (fecundity), but if that same individual also carries an allele that results in a fatal childhood disease, that fecundity phenotype will not be passed on to the next generation because the individual will not live to reach reproductive age. Natural selection acts at the level of the individual; it selects for individuals with greater contributions to the gene pool of the next generation, known as an organism's **evolutionary (Darwinian) fitness**.

Fitness is often quantifiable and is measured by scientists in the field. However, it is not the absolute fitness of an individual that counts, but rather how it compares to the other organisms in the population. This concept, called **relative fitness**, allows researchers to determine which individuals are contributing additional offspring to the next generation, and thus, how the population might evolve.

There are several ways selection can affect population variation: stabilizing selection, directional selection, diversifying selection, frequency-dependent selection, and sexual selection. As natural selection influences the allele frequencies in a population, individuals can either become more or less genetically similar and the phenotypes displayed can become more similar or more disparate.

## Stabilizing Selection



If natural selection favors an average phenotype, selecting against extreme variation, the population will undergo **stabilizing selection** ([\[link\]](#)). In a population of mice that live in the woods, for example, natural selection is likely to favor individuals that best blend in with the forest floor and are less likely to be spotted by predators. Assuming the ground is a fairly consistent shade of brown, those mice whose fur is most closely matched to that color will be most likely to survive and reproduce, passing on their genes for their brown coat. Mice that carry alleles that make them a bit lighter or a bit darker will stand out against the ground and be more likely to fall victim to predation. As a result of this selection, the population's genetic variance will decrease.

## Directional Selection

When the environment changes, populations will often undergo **directional selection** ([\[link\]](#)), which selects for phenotypes at one end of the spectrum of existing variation. A classic example of this type of selection is the evolution of the peppered moth in eighteenth- and nineteenth-century England. Prior to the Industrial Revolution, the moths were predominately light in color, which allowed them to blend in with the light-colored trees and lichens in their environment. But as soot began spewing from factories, the trees became darkened, and the light-colored moths became easier for predatory birds to spot. Over time, the frequency of the melanic form of the moth increased because they had a higher survival rate in habitats affected by air pollution because their darker coloration blended with the sooty trees. Similarly, the hypothetical mouse population may evolve to take on a different coloration if something were to cause the forest floor where they live to change color. The result of this type of selection is a shift in the population's genetic variance toward the new, fit phenotype.

### Note:

Link to Learning



In science, sometimes things are believed to be true, and then new information comes to light that changes our understanding. The story of the peppered moth is an example: the facts behind the selection toward darker moths have recently been called into question. Read this [article](#) to learn more.

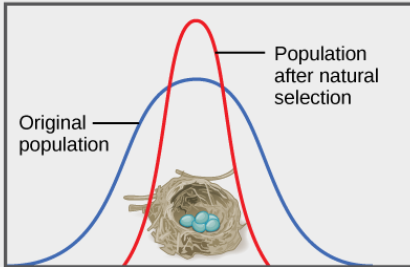
## Diversifying Selection

Sometimes two or more distinct phenotypes can each have their advantages and be selected for by natural selection, while the intermediate phenotypes are, on average, less fit. Known as **diversifying selection** ([link](#)), this is seen in many populations of animals that have multiple male forms. Large, dominant alpha males obtain mates by brute force, while small males can sneak in for furtive copulations with the females in an alpha male's territory. In this case, both the alpha males and the "sneaking" males will be selected for, but medium-sized males, which can't overtake the alpha males and are too big to sneak copulations, are selected against. Diversifying selection can also occur when environmental changes favor individuals on either end of the phenotypic spectrum. Imagine a population of mice living at the beach where there is light-colored sand interspersed with patches of tall grass. In this scenario, light-colored mice that blend in with the sand would be favored, as well as dark-colored mice that can hide in the grass. Medium-colored mice, on the other hand, would not blend in with either the grass or the sand, and would thus be more likely to be eaten by predators. The result of this type of selection is increased genetic variance as the population becomes more diverse.

## Note:

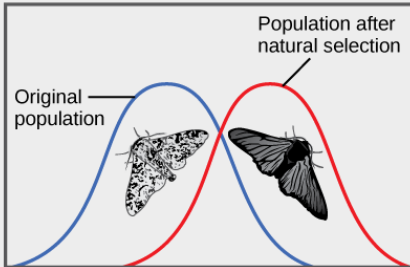
### Art Connection

(a) Stabilizing selection



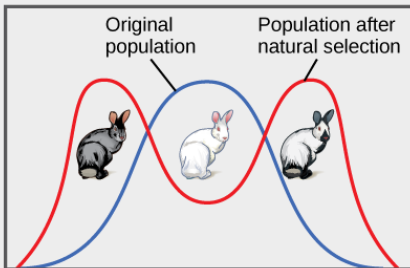
Robins typically lay four eggs, an example of stabilizing selection. Larger clutches may result in malnourished chicks, while smaller clutches may result in no viable offspring.

(b) Directional selection



Light-colored peppered moths are better camouflaged against a pristine environment; likewise, dark-colored peppered moths are better camouflaged against a sooty environment. Thus, as the Industrial Revolution progressed in nineteenth-century England, the color of the moth population shifted from light to dark, an example of directional selection.

(c) Diversifying selection



In a hypothetical population, gray and Himalayan (gray and white) rabbits are better able to blend with a rocky environment than white rabbits, resulting in diversifying selection.

Different types of natural selection can impact the distribution of phenotypes within a population. In (a) stabilizing selection, an average phenotype is favored. In (b) directional selection, a change in the environment shifts the spectrum of phenotypes observed. In (c) diversifying selection, two or more extreme phenotypes are selected for, while the average phenotype is selected against.

In recent years, factories have become cleaner, and less soot is released into the environment. What impact do you think this has had on the distribution of moth color in the population?

## Frequency-dependent Selection

Another type of selection, called **frequency-dependent selection**, favors phenotypes that are either common (positive frequency-dependent selection) or rare (negative frequency-dependent selection). An interesting example of this type of selection is seen in a unique group of lizards of the Pacific Northwest. Male common side-blotched lizards come in three throat-color patterns: orange, blue, and yellow. Each of these forms has a different reproductive strategy: orange males are the strongest and can fight other males for access to their females; blue males are medium-sized and form strong pair bonds with their mates; and yellow males ([link](#)) are the smallest, and look a bit like females, which allows them to sneak copulations. Like a game of rock-paper-scissors, orange beats blue, blue beats yellow, and yellow beats orange in the competition for females. That is, the big, strong orange males can fight off the blue males to mate with the blue's pair-bonded females, the blue males are successful at guarding their mates against yellow sneaker males, and the yellow males can sneak copulations from the potential mates of the large, polygynous orange males.



A yellow-throated side-blotched lizard is smaller than either the blue-throated or orange-throated males and appears a bit like the females of the species, allowing it to sneak copulations. (credit: “tinyfroglet”/Flickr)

In this scenario, orange males will be favored by natural selection when the population is dominated by blue males, blue males will thrive when the population is mostly yellow males, and yellow males will be selected for when orange males are the most populous. As a result, populations of side-blotched lizards cycle in the distribution of these phenotypes—in one generation, orange might be predominant, and then yellow males will begin to rise in frequency. Once yellow males make up a majority of the population, blue males will be selected for. Finally, when blue males become common, orange males will once again be favored.

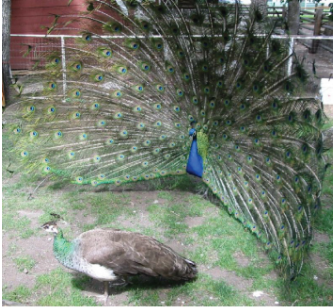
Negative frequency-dependent selection serves to increase the population’s genetic variance by selecting for rare phenotypes, whereas positive

frequency-dependent selection usually decreases genetic variance by selecting for common phenotypes.

## Sexual Selection

Males and females of certain species are often quite different from one another in ways beyond the reproductive organs. Males are often larger, for example, and display many elaborate colors and adornments, like the peacock's tail, while females tend to be smaller and duller in decoration. Such differences are known as **sexual dimorphisms** ([link](#)), which arise from the fact that in many populations, particularly animal populations, there is more variance in the reproductive success of the males than there is of the females. That is, some males—often the bigger, stronger, or more decorated males—get the vast majority of the total matings, while others receive none. This can occur because the males are better at fighting off other males, or because females will choose to mate with the bigger or more decorated males. In either case, this variation in reproductive success generates a strong selection pressure among males to get those matings, resulting in the evolution of bigger body size and elaborate ornaments to get the females' attention. Females, on the other hand, tend to get a handful of selected matings; therefore, they are more likely to select more desirable males.

Sexual dimorphism varies widely among species, of course, and some species are even sex-role reversed. In such cases, females tend to have a greater variance in their reproductive success than males and are correspondingly selected for the bigger body size and elaborate traits usually characteristic of males.



(a)



(b)



(c)

Sexual dimorphism is observed in (a) peacocks and peahens, (b) *Argiope appensa* spiders (the female spider is the large one), and in (c) wood ducks. (credit “spiders”: modification of work by “Sanba38”/Wikimedia Commons; credit “duck”: modification of work by Kevin Cole)

The selection pressures on males and females to obtain matings is known as sexual selection; it can result in the development of secondary sexual characteristics that do not benefit the individual’s likelihood of survival but help to maximize its reproductive success. Sexual selection can be so strong that it selects for traits that are actually detrimental to the individual’s survival. Think, once again, about the peacock’s tail. While it is beautiful and the male with the largest, most colorful tail is more likely to win the female, it is not the most practical appendage. In addition to being more visible to predators, it makes the males slower in their attempted escapes. There is some evidence that this risk, in fact, is why females like the big tails in the first place. The speculation is that large tails carry risk, and only the best males survive that risk: the bigger the tail, the more fit the male. This idea is known as the **handicap principle**.

The **good genes hypothesis** states that males develop these impressive ornaments to show off their efficient metabolism or their ability to fight disease. Females then choose males with the most impressive traits because it signals their genetic superiority, which they will then pass on to their offspring. Though it might be argued that females should not be picky because it will likely reduce their number of offspring, if better males father



more fit offspring, it may be beneficial. Fewer, healthier offspring may increase the chances of survival more than many, weaker offspring.

**Note:**

Link to Learning



In 1915, biologist Ronald Fisher proposed another model of sexual selection: the [Fisherian runaway model](#), which suggests that selection of certain traits is a result of sexual preference.

In both the handicap principle and the good genes hypothesis, the trait is said to be an **honest signal** of the males' quality, thus giving females a way to find the fittest mates—males that will pass the best genes to their offspring.

## No Perfect Organism

Natural selection is a driving force in evolution and can generate populations that are better adapted to survive and successfully reproduce in their environments. But natural selection cannot produce the perfect organism. Natural selection can only select on existing variation in the population; it does not create anything from scratch. Thus, it is limited by a population's existing genetic variance and whatever new alleles arise through mutation and gene flow.

Natural selection is also limited because it works at the level of individuals, not alleles, and some alleles are linked due to their physical proximity in the



genome, making them more likely to be passed on together (linkage disequilibrium). Any given individual may carry some beneficial alleles and some unfavorable alleles. It is the net effect of these alleles, or the organism's fitness, upon which natural selection can act. As a result, good alleles can be lost if they are carried by individuals that also have several overwhelmingly bad alleles; likewise, bad alleles can be kept if they are carried by individuals that have enough good alleles to result in an overall fitness benefit.

Furthermore, natural selection can be constrained by the relationships between different polymorphisms. One morph may confer a higher fitness than another, but may not increase in frequency due to the fact that going from the less beneficial to the more beneficial trait would require going through a less beneficial phenotype. Think back to the mice that live at the beach. Some are light-colored and blend in with the sand, while others are dark and blend in with the patches of grass. The dark-colored mice may be, overall, more fit than the light-colored mice, and at first glance, one might expect the light-colored mice be selected for a darker coloration. But remember that the intermediate phenotype, a medium-colored coat, is very bad for the mice—they cannot blend in with either the sand or the grass and are more likely to be eaten by predators. As a result, the light-colored mice would not be selected for a dark coloration because those individuals that began moving in that direction (began being selected for a darker coat) would be less fit than those that stayed light.

Finally, it is important to understand that not all evolution is adaptive. While natural selection selects the fittest individuals and often results in a more fit population overall, other forces of evolution, including genetic drift and gene flow, often do the opposite: introducing deleterious alleles to the population's gene pool. Evolution has no purpose—it is not changing a population into a preconceived ideal. It is simply the sum of the various forces described in this chapter and how they influence the genetic and phenotypic variance of a population.

## **Section Summary**

Because natural selection acts to increase the frequency of beneficial alleles and traits while decreasing the frequency of deleterious qualities, it is adaptive evolution. Natural selection acts at the level of the individual, selecting for those that have a higher overall fitness compared to the rest of the population. If the fit phenotypes are those that are similar, natural selection will result in stabilizing selection, and an overall decrease in the population's variation. Directional selection works to shift a population's variance toward a new, fit phenotype, as environmental conditions change. In contrast, diversifying selection results in increased genetic variance by selecting for two or more distinct phenotypes.

Other types of selection include frequency-dependent selection, in which individuals with either common (positive frequency-dependent selection) or rare (negative frequency-dependent selection) are selected for. Finally, sexual selection results from the fact that one sex has more variance in the reproductive success than the other. As a result, males and females experience different selective pressures, which can often lead to the evolution of phenotypic differences, or sexual dimorphisms, between the two.

## Art Connection

### Exercise:

#### Problem:

[\[link\]](#) In recent years, factories have become cleaner, and less soot is released into the environment. What impact do you think this has had on the distribution of moth color in the population?

---

#### Solution:

[\[link\]](#) Moths have shifted to a lighter color.

## Review Questions

### Exercise:

**Problem:**

Which type of selection results in greater genetic variance in a population?

- a. stabilizing selection
- b. directional selection
- c. diversifying selection
- d. positive frequency-dependent selection

---

**Solution:**

C

**Exercise:**

**Problem:**

When males and females of a population look or act differently, it is referred to as \_\_\_\_\_.

- a. sexual dimorphism
- b. sexual selection
- c. diversifying selection
- d. a cline

---

**Solution:**

A

**Exercise:**

**Problem:** The good genes hypothesis is a theory that explains what?

- a. why more fit individuals are more likely to have more offspring
- b. why alleles that confer beneficial traits or behaviors are selected for by natural selection
- c. why some deleterious mutations are maintained in the population

d. why individuals of one sex develop impressive ornamental traits

---

**Solution:**

D

## Free Response

**Exercise:**

**Problem:**

Give an example of a trait that may have evolved as a result of the handicap principle and explain your reasoning.

---

**Solution:**

The peacock's tail is a good example of the handicap principle. The tail, which makes the males more visible to predators and less able to escape, is clearly a disadvantage to the bird's survival. But because it is a disadvantage, only the most fit males should be able to survive with it. Thus, the tail serves as an honest signal of quality to the females of the population; therefore, the male will earn more matings and greater reproductive success.

**Exercise:**

**Problem:**

List the ways in which evolution can affect population variation and describe how they influence allele frequencies.

---

**Solution:**

There are several ways evolution can affect population variation: stabilizing selection, directional selection, diversifying selection, frequency-dependent selection, and sexual selection. As these influence the allele frequencies in a population, individuals can either

become more or less related, and the phenotypes displayed can become more similar or more disparate.

## **Glossary**

adaptive evolution

increase in frequency of beneficial alleles and decrease in deleterious alleles due to selection

directional selection

selection that favors phenotypes at one end of the spectrum of existing variation

diversifying selection

selection that favors two or more distinct phenotypes

evolutionary fitness

(also, Darwinian fitness) individual's ability to survive and reproduce

frequency-dependent selection

selection that favors phenotypes that are either common (positive frequency-dependent selection) or rare (negative frequency-dependent selection)

good genes hypothesis

theory of sexual selection that argues individuals develop impressive ornaments to show off their efficient metabolism or ability to fight disease

handicap principle

theory of sexual selection that argues only the fittest individuals can afford costly traits

honest signal

trait that gives a truthful impression of an individual's fitness

relative fitness

individual's ability to survive and reproduce relative to the rest of the population

sexual dimorphism

phenotypic difference between the males and females of a population

stabilizing selection

selection that favors average phenotypes

## Introduction

class="introduction"

The life of a  
bee is very  
different  
from the life  
of a flower,  
but the two  
organisms  
are related.

Both are  
members  
the domain  
Eukarya and  
have cells  
containing  
many  
similar  
organelles,  
genes, and  
proteins.

(credit:  
modificatio  
n of work  
by John  
Beetham)



This bee and *Echinacea* flower ([\[link\]](#)) could not look more different, yet they are related, as are all living organisms on Earth. By following pathways of similarities and changes—both visible and genetic—scientists seek to map the evolutionary past of how life developed from single-celled organisms to the tremendous collection of creatures that have germinated, crawled, floated, swam, flown, and walked on this planet.



## Organizing Life on Earth

By the end of this section, you will be able to:

- Discuss the need for a comprehensive classification system
- List the different levels of the taxonomic classification system
- Describe how systematics and taxonomy relate to phylogeny
- Discuss the components and purpose of a phylogenetic tree

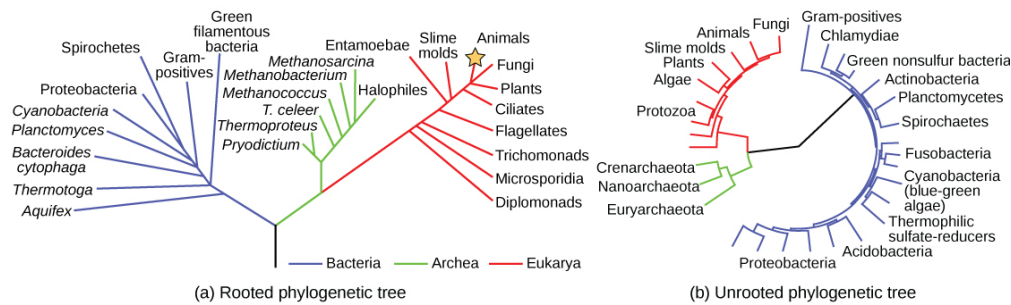
In scientific terms, the evolutionary history and relationship of an organism or group of organisms is called its **phylogeny**. A phylogeny describes the relationships of an organism, such as from which organisms it is thought to have evolved, to which species it is most closely related, and so forth.

Phylogenetic relationships provide information on shared ancestry but not necessarily on how organisms are similar or different.

## Phylogenetic Trees

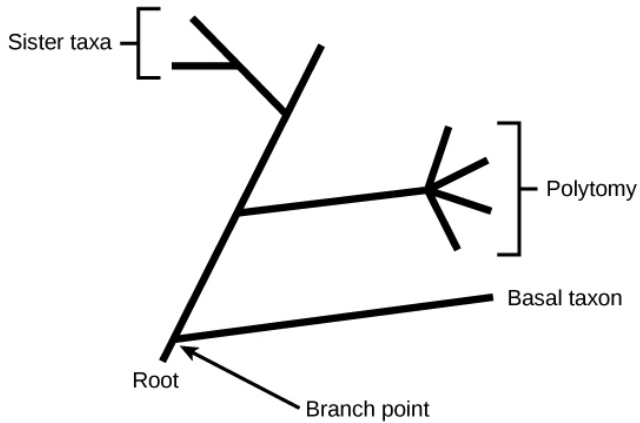
Scientists use a tool called a phylogenetic tree to show the evolutionary pathways and connections among organisms. A **phylogenetic tree** is a diagram used to reflect evolutionary relationships among organisms or groups of organisms. Scientists consider phylogenetic trees to be a hypothesis of the evolutionary past since one cannot go back to confirm the proposed relationships. In other words, a “tree of life” can be constructed to illustrate when different organisms evolved and to show the relationships among different organisms ([\[link\]](#)).

Unlike a taxonomic classification diagram, a phylogenetic tree can be read like a map of evolutionary history. Many phylogenetic trees have a single lineage at the base representing a common ancestor. Scientists call such trees **rooted**, which means there is a single ancestral lineage (typically drawn from the bottom or left) to which all organisms represented in the diagram relate. Notice in the rooted phylogenetic tree that the three domains — Bacteria, Archaea, and Eukarya—diverge from a single point and branch off. The small branch that plants and animals (including humans) occupy in this diagram shows how recent and miniscule these groups are compared with other organisms. Unrooted trees don’t show a common ancestor but do show relationships among species.



Both of these phylogenetic trees shows the relationship of the three domains of life—Bacteria, Archaea, and Eukarya—but the (a) rooted tree attempts to identify when various species diverged from a common ancestor while the (b) unrooted tree does not. (credit a: modification of work by Eric Gaba)

In a rooted tree, the branching indicates evolutionary relationships ([link](#)). The point where a split occurs, called a **branch point**, represents where a single lineage evolved into a distinct new one. A lineage that evolved early from the root and remains unbranched is called **basal taxon**. When two lineages stem from the same branch point, they are called **sister taxa**. A branch with more than two lineages is called a **polytomy** and serves to illustrate where scientists have not definitively determined all of the relationships. It is important to note that although sister taxa and polytomy do share an ancestor, it does not mean that the groups of organisms split or evolved from each other. Organisms in two taxa may have split apart at a specific branch point, but neither taxa gave rise to the other.



The root of a phylogenetic tree indicates that an ancestral lineage gave rise to all organisms on the tree. A branch point indicates where two lineages diverged. A lineage that evolved early and remains unbranched is a basal taxon. When two lineages stem from the same branch point, they are sister taxa. A branch with more than two lineages is a polytomy.

The diagrams above can serve as a pathway to understanding evolutionary history. The pathway can be traced from the origin of life to any individual species by navigating through the evolutionary branches between the two points. Also, by starting with a single species and tracing back towards the "trunk" of the tree, one can discover that species' ancestors, as well as where lineages share a common ancestry. In addition, the tree can be used to study entire groups of organisms.

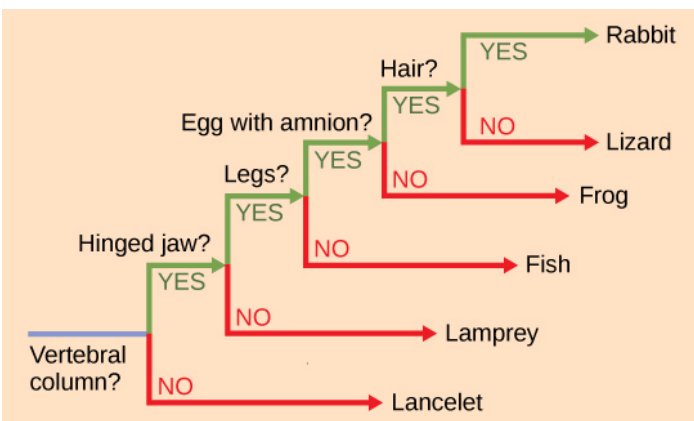
Another point to mention on phylogenetic tree structure is that rotation at branch points does not change the information. For example, if a branch point was rotated and the taxon order changed, this would not alter the

information because the evolution of each taxon from the branch point was independent of the other.

Many disciplines within the study of biology contribute to understanding how past and present life evolved over time; these disciplines together contribute to building, updating, and maintaining the “tree of life.” Information is used to organize and classify organisms based on evolutionary relationships in a scientific field called **systematics**. Data may be collected from fossils, from studying the structure of body parts or molecules used by an organism, and by DNA analysis. By combining data from many sources, scientists can put together the phylogeny of an organism; since phylogenetic trees are hypotheses, they will continue to change as new types of life are discovered and new information is learned.

## Limitations of Phylogenetic Trees

It may be easy to assume that more closely related organisms look more alike, and while this is often the case, it is not always true. If two closely related lineages evolved under significantly varied surroundings or after the evolution of a major new adaptation, it is possible for the two groups to appear more different than other groups that are not as closely related. For example, the phylogenetic tree in [\[link\]](#) shows that lizards and rabbits both have amniotic eggs, whereas frogs do not; yet lizards and frogs appear more similar than lizards and rabbits.



This ladder-like phylogenetic tree of vertebrates is rooted by an organism that lacked a vertebral column. At each branch point, organisms with different characters are placed in different groups based on the characteristics they share.

Another aspect of phylogenetic trees is that, unless otherwise indicated, the branches do not account for length of time, only the evolutionary order. In other words, the length of a branch does not typically mean more time passed, nor does a short branch mean less time passed— unless specified on the diagram. For example, in [\[link\]](#), the tree does not indicate how much time passed between the evolution of amniotic eggs and hair. What the tree does show is the order in which things took place. Again using [\[link\]](#), the tree shows that the oldest trait is the vertebral column, followed by hinged jaws, and so forth. Remember that any phylogenetic tree is a part of the greater whole, and like a real tree, it does not grow in only one direction after a new branch develops. So, for the organisms in [\[link\]](#), just because a vertebral column evolved does not mean that invertebrate evolution ceased, it only means that a new branch formed. Also, groups that are not closely related, but evolve under similar conditions, may appear more phenotypically similar to each other than to a close relative.

**Note:****Link to Learning**

Head to this [website](#) to see interactive exercises that allow you to explore the evolutionary relationships among species.

## The Levels of Classification

**Taxonomy** (which literally means “arrangement law”) is the science of classifying organisms to construct internationally shared classification systems with each organism placed into more and more inclusive groupings. Think about how a grocery store is organized. One large space is divided into departments, such as produce, dairy, and meats. Then each department further divides into aisles, then each aisle into categories and brands, and then finally a single product. This organization from larger to smaller, more specific categories is called a hierarchical system.

The taxonomic classification system (also called the Linnaean system after its inventor, Carl Linnaeus, a Swedish botanist, zoologist, and physician) uses a hierarchical model. Moving from the point of origin, the groups become more specific, until one branch ends as a single species. For example, after the common beginning of all life, scientists divide organisms into three large categories called a domain: Bacteria, Archaea, and Eukarya. Within each domain is a second category called a **kingdom**. After kingdoms, the subsequent categories of increasing specificity are: **phylum**, **class**, **order**, **family**, **genus**, and **species** ([link](#)).



**Subspecies:** *Canus lupus familiaris*

**Species:** *Canis lupus*

**Genus:** *Canis*

**Family:** Canidae

**Order:** Carnivora

**Class:** Mammalia

**Phylum:** Chordata

**Kingdom:** Animalia

**Domain:** Eukarya

The taxonomic classification system uses a hierarchical model to organize living organisms into increasingly specific categories. The common dog, *Canis lupus familiaris*, is a subspecies of *Canis lupus*, which also includes the wolf and dingo. (credit “dog”:

modification of work by Janneke  
Vreugdenhil)














































The kingdom Animalia stems from the Eukarya domain. For the common dog, the classification levels would be as shown in [\[link\]](#). Therefore, the full name of an organism technically has eight terms. For the dog, it is: Eukarya, Animalia, Chordata, Mammalia, Carnivora, Canidae, *Canis*, and *lupus*. Notice that each name is capitalized except for species, and the genus and species names are italicized. Scientists generally refer to an organism only by its genus and species, which is its two-word scientific name, in what is called **binomial nomenclature**. Therefore, the scientific name of the dog is *Canis lupus*. The name at each level is also called a **taxon**. In other words, dogs are in order Carnivora. Carnivora is the name of the taxon at the order level; Canidae is the taxon at the family level, and so forth. Organisms also have a common name that people typically use, in this case, dog. Note that the dog is additionally a subspecies: the “*familiaris*” in *Canis lupus familiaris*. Subspecies are members of the same species that are capable of mating and reproducing viable offspring, but they are considered separate subspecies due to geographic or behavioral isolation or other factors.

[\[link\]](#) shows how the levels move toward specificity with other organisms. Notice how the dog shares a domain with the widest diversity of organisms, including plants and butterflies. At each sublevel, the organisms become more similar because they are more closely related. Historically, scientists classified organisms using characteristics, but as DNA technology developed, more precise phylogenies have been determined.

**Note:**

Art Connection



<b>Subspecies:</b> <i>Canis lupus familiaris</i>	 Dog																																																																																									
<b>Species:</b> <i>Canis lupus</i>	 Wolf										 Dog																																																																															
<b>Genus:</b> <i>Canis</i>	 Jackal										 Wolf										 Dog																																																																					
<b>Family:</b> Canidae	 Fox										 Jackal										 Wolf										 Dog																																																											
<b>Order:</b> Carnivora	 Cat										 Fox										 Jackal										 Wolf										 Dog																																																	
<b>Class:</b> Mammalia	 Rabbit										 Cat										 Fox										 Jackal										 Wolf										 Dog																																							
<b>Phylum:</b> Chordata	 Fish										 Rabbit										 Cat										 Fox										 Jackal										 Wolf										 Dog																													
<b>Kingdom:</b> Animalia	 Insect										 Fish										 Rabbit										 Cat										 Fox										 Jackal										 Wolf										 Dog																			
<b>Domain:</b> Eukarya	 Plant										 Insect										 Fish										 Rabbit										 Cat										 Fox										 Jackal										 Wolf										 Dog									

At each sublevel in the taxonomic classification system, organisms become more similar. Dogs and wolves are the same species because they can breed and produce viable offspring, but they are different enough to be classified as different subspecies. (credit “plant”: modification of work by "berduchwal"/Flickr; credit “insect”: modification of work by Jon Sullivan; credit “fish”: modification of work by Christian Mehlführer; credit “rabbit”: modification of work by Aidan Wojtas; credit “cat”: modification of work by Jonathan Lidbeck; credit “fox”: modification of work by Kevin Bacher, NPS; credit “jackal”: modification of work by Thomas A. Hermann, NBII, USGS; credit “wolf”: modification

of work by Robert Dewar; credit “dog”: modification of work by "digital\_image\_fan"/Flickr)

At what levels are cats and dogs considered to be part of the same group?

**Note:**

Link to Learning



Visit this [website](#) to classify three organisms—bear, orchid, and sea cucumber—from kingdom to species. To launch the game, under Classifying Life, click the picture of the bear or the Launch Interactive button.

Recent genetic analysis and other advancements have found that some earlier phylogenetic classifications do not align with the evolutionary past; therefore, changes and updates must be made as new discoveries occur. Recall that phylogenetic trees are hypotheses and are modified as data becomes available. In addition, classification historically has focused on grouping organisms mainly by shared characteristics and does not necessarily illustrate how the various groups relate to each other from an evolutionary perspective. For example, despite the fact that a hippopotamus resembles a pig more than a whale, the hippopotamus may be the closest living relative of the whale.

## Section Summary

Scientists continually gain new information that helps understand the evolutionary history of life on Earth. Each group of organisms went through its own evolutionary journey, called its phylogeny. Each organism shares relatedness with others, and based on morphologic and genetic evidence, scientists attempt to map the evolutionary pathways of all life on Earth. Historically, organisms were organized into a taxonomic classification system. However, today many scientists build phylogenetic trees to illustrate evolutionary relationships.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) At what levels are cats and dogs considered to be part of the same group?

---

#### Solution:

[\[link\]](#) Cats and dogs are part of the same group at five levels: both are in the domain Eukarya, the kingdom Animalia, the phylum Chordata, the class Mammalia, and the order Carnivora.

## Review Questions

### Exercise:

**Problem:** What is used to determine phylogeny?

- a. mutations
  - b. DNA
  - c. evolutionary history
  - d. organisms on earth
- 

#### Solution:

C

**Exercise:**

**Problem:** What do scientists in the field of systematics accomplish?

- a. discover new fossil sites
  - b. organize and classify organisms
  - c. name new species
  - d. communicate among field biologists
- 

**Solution:**

B

**Exercise:**

**Problem:**

Which statement about the taxonomic classification system is correct?

- a. There are more domains than kingdoms.
  - b. Kingdoms are the top category of classification.
  - c. Classes are divisions of orders.
  - d. Subspecies are the most specific category of classification.
- 

**Solution:**

D

**Exercise:**

**Problem:**

On a phylogenetic tree, which term refers to lineages that diverged from the same place?

- a. sister taxa
- b. basal taxa

- c. rooted taxa
  - d. dichotomous taxa
- 

**Solution:**

A

## Free Response

**Exercise:**

**Problem:** How does a phylogenetic tree relate to the passing of time?

---

**Solution:**

The phylogenetic tree shows the order in which evolutionary events took place and in what order certain characteristics and organisms evolved in relation to others. It does not relate to time.

**Exercise:**

**Problem:**

Some organisms that appear very closely related on a phylogenetic tree may not actually be closely related. Why is this?

---

**Solution:**

In most cases, organisms that appear closely related actually are; however, there are cases where organisms evolved through convergence and appear closely related but are not.

**Exercise:**

**Problem:**

List the different levels of the taxonomic classification system.

---

**Solution:**

domain, kingdom, phylum, class, order, family, genus, species

**Glossary**

basal taxon

branch on a phylogenetic tree that has not diverged significantly from the root ancestor

binomial nomenclature

system of two-part scientific names for an organism, which includes genus and species names

branch point

node on a phylogenetic tree where a single lineage splits into distinct new ones

class

division of phylum in the taxonomic classification system

family

division of order in the taxonomic classification system

genus

division of family in the taxonomic classification system; the first part of the binomial scientific name

kingdom

division of domain in the taxonomic classification system

order

division of class in the taxonomic classification system

phylogenetic tree

diagram used to reflect the evolutionary relationships among organisms or groups of organisms

phylogeny

evolutionary history and relationship of an organism or group of organisms

phylum

(plural: phyla) division of kingdom in the taxonomic classification system

polytomy

branch on a phylogenetic tree with more than two groups or taxa

rooted

single ancestral lineage on a phylogenetic tree to which all organisms represented in the diagram relate

sister taxa

two lineages that diverged from the same branch point

systematics

field of organizing and classifying organisms based on evolutionary relationships

taxon

(plural: taxa) single level in the taxonomic classification system

taxonomy

science of classifying organisms

## Determining Evolutionary Relationships

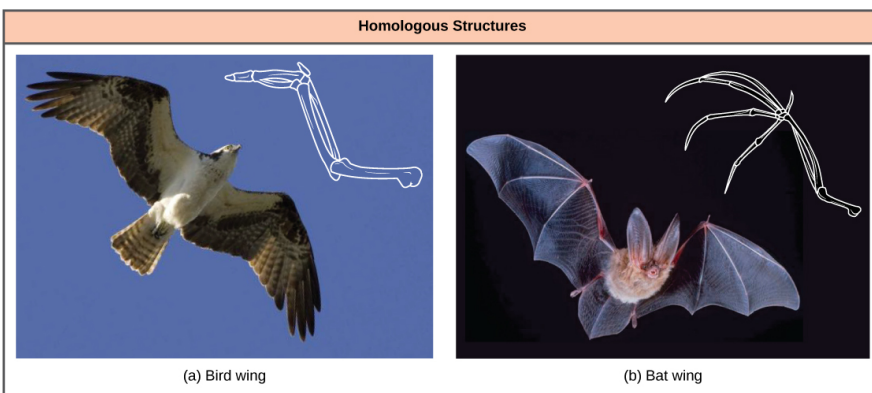
By the end of this section, you will be able to:

- Compare homologous and analogous traits
- Discuss the purpose of cladistics
- Describe maximum parsimony

Scientists must collect accurate information that allows them to make evolutionary connections among organisms. Similar to detective work, scientists must use evidence to uncover the facts. In the case of phylogeny, evolutionary investigations focus on two types of evidence: morphologic (form and function) and genetic.

## Two Options for Similarities

In general, organisms that share similar physical features and genomes tend to be more closely related than those that do not. Such features that overlap both morphologically (in form) and genetically are referred to as homologous structures; they stem from developmental similarities that are based on evolution. For example, the bones in the wings of bats and birds have homologous structures ([link](#)).



Bat and bird wings are homologous structures, indicating that bats and birds share a common evolutionary past. (credit a: modification of work



by Steve Hillebrand, USFWS; credit b:  
modification of work by U.S. DOI BLM)

Notice it is not simply a single bone, but rather a grouping of several bones arranged in a similar way. The more complex the feature, the more likely any kind of overlap is due to a common evolutionary past. Imagine two people from different countries both inventing a car with all the same parts and in exactly the same arrangement without any previous or shared knowledge. That outcome would be highly improbable. However, if two people both invented a hammer, it would be reasonable to conclude that both could have the original idea without the help of the other. The same relationship between complexity and shared evolutionary history is true for homologous structures in organisms.

## Misleading Appearances

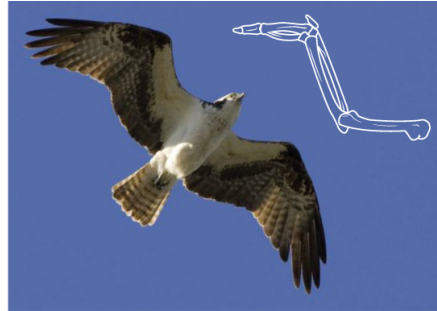
Some organisms may be very closely related, even though a minor genetic change caused a major morphological difference to make them look quite different. Similarly, unrelated organisms may be distantly related, but appear very much alike. This usually happens because both organisms were in common adaptations that evolved within similar environmental conditions. When similar characteristics occur because of environmental constraints and not due to a close evolutionary relationship, it is called an **analogy** or homoplasy. For example, insects use wings to fly like bats and birds, but the wing structure and embryonic origin is completely different. These are called analogous structures ([\[link\]](#)).

Similar traits can be either homologous or analogous. Homologous structures share a similar embryonic origin; analogous organs have a similar function. For example, the bones in the front flipper of a whale are homologous to the bones in the human arm. These structures are not analogous. The wings of a butterfly and the wings of a bird are analogous but not homologous. Some structures are both analogous and homologous: the wings of a bird and the wings of a bat are both homologous and

analogous. Scientists must determine which type of similarity a feature exhibits to decipher the phylogeny of the organisms being studied.



(a) Bat wing



(b) Bird wing



(c) Insect wing

The (c) wing of a honeybee is similar in shape to a (b) bird wing and (a) bat wing, and it serves the same function. However, the honeybee wing is not composed of bones and has a distinctly different structure and embryonic origin. These wing types (insect versus bat and bird) illustrate an analogy—similar structures that do not share an evolutionary history. (credit a: modification of work by Steve Hillebrand, USFWS; credit b: modification of work by U.S. DOI BLM; credit c: modification of work by Jon Sullivan)

**Note:**

Link to Learning



This [website](#) has several examples to show how appearances can be misleading in understanding the phylogenetic relationships of organisms.

## Molecular Comparisons

With the advancement of DNA technology, the area of **molecular systematics**, which describes the use of information on the molecular level including DNA analysis, has blossomed. New computer programs not only confirm many earlier classified organisms, but also uncover previously made errors. As with physical characteristics, even the DNA sequence can be tricky to read in some cases. For some situations, two very closely related organisms can appear unrelated if a mutation occurred that caused a shift in the genetic code. An insertion or deletion mutation would move each nucleotide base over one place, causing two similar codes to appear unrelated.

Sometimes two segments of DNA code in distantly related organisms randomly share a high percentage of bases in the same locations, causing these organisms to appear closely related when they are not. For both of these situations, computer technologies have been developed to help identify the actual relationships, and, ultimately, the coupled use of both morphologic and molecular information is more effective in determining phylogeny.

**Note:**

## Evolution Connection

### Why Does Phylogeny Matter?

Evolutionary biologists could list many reasons why understanding phylogeny is important to everyday life in human society. For botanists, phylogeny acts as a guide to discovering new plants that can be used to benefit people. Think of all the ways humans use plants—food, medicine, and clothing are a few examples. If a plant contains a compound that is effective in treating cancer, scientists might want to examine all of the relatives of that plant for other useful drugs.

A research team in China identified a segment of DNA thought to be common to some medicinal plants in the family Fabaceae (the legume family) and worked to identify which species had this segment ([link](#)). After testing plant species in this family, the team found a DNA marker (a known location on a chromosome that enabled them to identify the species) present. Then, using the DNA to uncover phylogenetic relationships, the team could identify whether a newly discovered plant was in this family and assess its potential medicinal properties.



*Dalbergia sissoo* (*D. sissoo*) is in the Fabaceae, or legume family.

Scientists found that *D. sissoo* shares a DNA marker with species within the Fabaceae family that have antifungal properties.

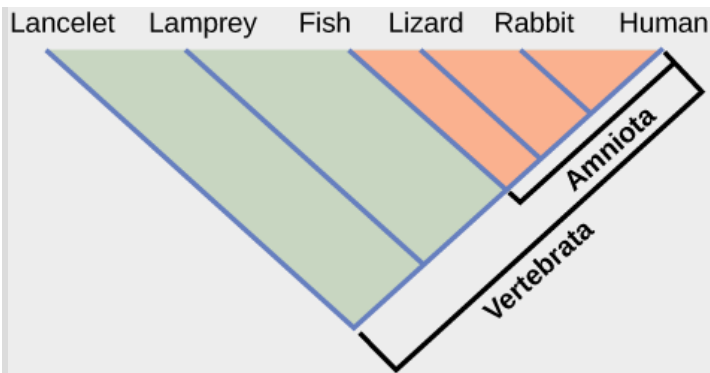
Subsequently, *D. sissoo* was shown to have fungicidal activity, supporting the idea that DNA markers can be used to screen for plants with potential medicinal properties.

## Building Phylogenetic Trees

How do scientists construct phylogenetic trees? After the homologous and analogous traits are sorted, scientists often organize the homologous traits using a system called **cladistics**. This system sorts organisms into clades: groups of organisms that descended from a single ancestor. For example, in [\[link\]](#), all of the organisms in the orange region evolved from a single ancestor that had amniotic eggs. Consequently, all of these organisms also have amniotic eggs and make a single clade, also called a **monophyletic group**. Clades must include all of the descendants from a branch point.

### Note:

Art Connection

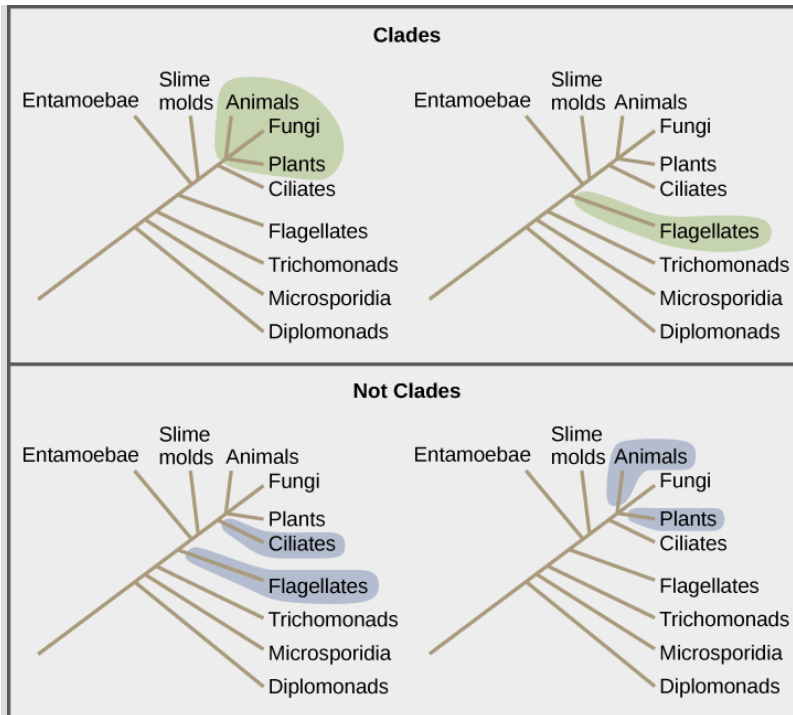


Lizards, rabbits, and humans all descend from a common ancestor that had an amniotic egg. Thus, lizards, rabbits, and humans all belong to the clade Amniota. Vertebrata is a larger clade that also includes fish and lamprey.

Which animals in this figure belong to a clade that includes animals with hair? Which evolved first, hair or the amniotic egg?

Clades can vary in size depending on which branch point is being referenced. The important factor is that all of the organisms in the clade or monophyletic group stem from a single point on the tree. This can be remembered because monophyletic breaks down into “mono,” meaning one, and “phyletic,” meaning evolutionary relationship. [\[link\]](#) shows various examples of clades. Notice how each clade comes from a single point, whereas the non-clade groups show branches that do not share a single point.

**Note:**  
Art Connection



All the organisms within a clade stem from a single point on the tree. A clade may contain multiple groups, as in the case of animals, fungi and plants, or a single group, as in the case of flagellates. Groups that diverge at a different branch point, or that do not include all groups in a single branch point, are not considered clades.

What is the largest clade in this diagram?

## Shared Characteristics

Organisms evolve from common ancestors and then diversify. Scientists use the phrase “descent with modification” because even though related organisms have many of the same characteristics and genetic codes, changes occur. This pattern repeats over and over as one goes through the phylogenetic tree of life:

1. A change in the genetic makeup of an organism leads to a new trait which becomes prevalent in the group.
2. Many organisms descend from this point and have this trait.
3. New variations continue to arise: some are adaptive and persist, leading to new traits.
4. With new traits, a new branch point is determined (go back to step 1 and repeat).

If a characteristic is found in the ancestor of a group, it is considered a **shared ancestral character** because all of the organisms in the taxon or clade have that trait. The vertebrate in [\[link\]](#) is a shared ancestral character. Now consider the amniotic egg characteristic in the same figure. Only some of the organisms in [\[link\]](#) have this trait, and to those that do, it is called a **shared derived character** because this trait derived at some point but does not include all of the ancestors in the tree.

The tricky aspect to shared ancestral and shared derived characters is the fact that these terms are relative. The same trait can be considered one or the other depending on the particular diagram being used. Returning to [\[link\]](#), note that the amniotic egg is a shared ancestral character for the Amniota clade, while having hair is a shared derived character for some organisms in this group. These terms help scientists distinguish between clades in the building of phylogenetic trees.

### Choosing the Right Relationships

Imagine being the person responsible for organizing all of the items in a department store properly—an overwhelming task. Organizing the evolutionary relationships of all life on Earth proves much more difficult: scientists must span enormous blocks of time and work with information from long-extinct organisms. Trying to decipher the proper connections, especially given the presence of homologies and analogies, makes the task of building an accurate tree of life extraordinarily difficult. Add to that the advancement of DNA technology, which now provides large quantities of genetic sequences to be used and analyzed. Taxonomy is a subjective



discipline: many organisms have more than one connection to each other, so each taxonomist will decide the order of connections.

To aid in the tremendous task of describing phylogenies accurately, scientists often use a concept called **maximum parsimony**, which means that events occurred in the simplest, most obvious way. For example, if a group of people entered a forest preserve to go hiking, based on the principle of maximum parsimony, one could predict that most of the people would hike on established trails rather than forge new ones.

For scientists deciphering evolutionary pathways, the same idea is used: the pathway of evolution probably includes the fewest major events that coincide with the evidence at hand. Starting with all of the homologous traits in a group of organisms, scientists look for the most obvious and simple order of evolutionary events that led to the occurrence of those traits.

**Note:**

Link to Learning



Head to this [website](#) to learn how maximum parsimony is used to create phylogenetic trees.

These tools and concepts are only a few of the strategies scientists use to tackle the task of revealing the evolutionary history of life on Earth.

Recently, newer technologies have uncovered surprising discoveries with unexpected relationships, such as the fact that people seem to be more closely related to fungi than fungi are to plants. Sound unbelievable? As the information about DNA sequences grows, scientists will become closer to mapping the evolutionary history of all life on Earth.

## Section Summary

To build phylogenetic trees, scientists must collect accurate information that allows them to make evolutionary connections between organisms. Using morphologic and molecular data, scientists work to identify homologous characteristics and genes. Similarities between organisms can stem either from shared evolutionary history (homologies) or from separate evolutionary paths (analogies). Newer technologies can be used to help distinguish homologies from analogies. After homologous information is identified, scientists use cladistics to organize these events as a means to determine an evolutionary timeline. Scientists apply the concept of maximum parsimony, which states that the order of events probably occurred in the most obvious and simple way with the least amount of steps. For evolutionary events, this would be the path with the least number of major divergences that correlate with the evidence.

## Art Connections

### Exercise:

#### Problem:

[\[link\]](#) Which animals in this figure belong to a clade that includes animals with hair? Which evolved first, hair or the amniotic egg?

---

#### Solution:

[\[link\]](#) Rabbits and humans belong in the clade that includes animals with hair. The amniotic egg evolved before hair because the Amniota clade is larger than the clade that encompasses animals with hair.

### Exercise:

**Problem:** [\[link\]](#) What is the largest clade in this diagram?

---

#### Solution:

[\[link\]](#) The largest clade encompasses the entire tree.

## Review Questions

### Exercise:

**Problem:** Which statement about analogies is correct?

- a. They occur only as errors.
- b. They are synonymous with homologous traits.
- c. They are derived by similar environmental constraints.
- d. They are a form of mutation.

---

**Solution:**

C

### Exercise:

**Problem:** What do scientists use to apply cladistics?

- a. homologous traits
- b. homoplasies
- c. analogous traits
- d. monophyletic groups

---

**Solution:**

A

### Exercise:

**Problem:**

What is true about organisms that are a part of the same clade?

- a. They all share the same basic characteristics.
- b. They evolved from a shared ancestor.

- c. They usually fall into the same classification taxa.
- d. They have identical phylogenies.

---

**Solution:**

B

**Exercise:**

**Problem:**

Why do scientists apply the concept of maximum parsimony?

- a. to decipher accurate phylogenies
- b. to eliminate analogous traits
- c. to identify mutations in DNA codes
- d. to locate homoplasies

---

**Solution:**

A

**Free Response**

**Exercise:**

**Problem:**

Dolphins and fish have similar body shapes. Is this feature more likely a homologous or analogous trait?

---

**Solution:**

Dolphins are mammals and fish are not, which means that their evolutionary paths (phylogenies) are quite separate. Dolphins probably adapted to have a similar body plan after returning to an aquatic lifestyle, and, therefore, this trait is probably analogous.

**Exercise:****Problem:**

Why is it so important for scientists to distinguish between homologous and analogous characteristics before building phylogenetic trees?

---

**Solution:**

Phylogenetic trees are based on evolutionary connections. If an analogous similarity were used on a tree, this would be erroneous and, furthermore, would cause the subsequent branches to be inaccurate.

**Exercise:**

**Problem:** Describe maximum parsimony.

---

**Solution:**

Maximum parsimony hypothesizes that events occurred in the simplest, most obvious way, and the pathway of evolution probably includes the fewest major events that coincide with the evidence at hand.

**Glossary**

analogy

(also, homoplasy) characteristic that is similar between organisms by convergent evolution, not due to the same evolutionary path

cladistics

system used to organize homologous traits to describe phylogenies

maximum parsimony

applying the simplest, most obvious way with the least number of steps

molecular systematics

technique using molecular evidence to identify phylogenetic relationships

monophyletic group

(also, clade) organisms that share a single ancestor

shared ancestral character

describes a characteristic on a phylogenetic tree that is shared by all organisms on the tree

shared derived character

describes a characteristic on a phylogenetic tree that is shared only by a certain clade of organisms

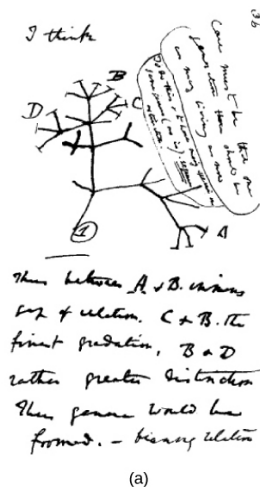
## Perspectives on the Phylogenetic Tree

By the end of this section, you will be able to:

- Describe horizontal gene transfer
- Illustrate how prokaryotes and eukaryotes transfer genes horizontally
- Identify the web and ring models of phylogenetic relationships and describe how they differ from the original phylogenetic tree concept

The concepts of phylogenetic modeling are constantly changing. It is one of the most dynamic fields of study in all of biology. Over the last several decades, new research has challenged scientists' ideas about how organisms are related. New models of these relationships have been proposed for consideration by the scientific community.

Many phylogenetic trees have been shown as models of the evolutionary relationship among species. Phylogenetic trees originated with Charles Darwin, who sketched the first phylogenetic tree in 1837 ([link](#)a), which served as a pattern for subsequent studies for more than a century. The concept of a phylogenetic tree with a single trunk representing a common ancestor, with the branches representing the divergence of species from this ancestor, fits well with the structure of many common trees, such as the oak ([link](#)b). However, evidence from modern DNA sequence analysis and newly developed computer algorithms has caused skepticism about the validity of the standard tree model in the scientific community.



The (a) concept of the “tree of life” goes back to an 1837 sketch by Charles Darwin. Like an (b) oak tree, the “tree of life” has a single trunk and many branches. (credit b:

modification of work by "Amada44"/Wikimedia Commons)

## **Limitations to the Classic Model**

Classical thinking about prokaryotic evolution, included in the classic tree model, is that species evolve clonally. That is, they produce offspring themselves with only random mutations causing the descent into the variety of modern-day and extinct species known to science. This view is somewhat complicated in eukaryotes that reproduce sexually, but the laws of Mendelian genetics explain the variation in offspring, again, to be a result of a mutation within the species. The concept of genes being transferred between unrelated species was not considered as a possibility until relatively recently. Horizontal gene transfer (HGT), also known as lateral gene transfer, is the transfer of genes between unrelated species. HGT has been shown to be an ever-present phenomenon, with many evolutionists postulating a major role for this process in evolution, thus complicating the simple tree model. Genes have been shown to be passed between species which are only distantly related using standard phylogeny, thus adding a layer of complexity to the understanding of phylogenetic relationships.

The various ways that HGT occurs in prokaryotes is important to understanding phylogenies. Although at present HGT is not viewed as important to eukaryotic evolution, HGT does occur in this domain as well. Finally, as an example of the ultimate gene transfer, theories of genome fusion between symbiotic or endosymbiotic organisms have been proposed to explain an event of great importance—the evolution of the first eukaryotic cell, without which humans could not have come into existence.

## **Horizontal Gene Transfer**

**Horizontal gene transfer (HGT)** is the introduction of genetic material from one species to another species by mechanisms other than the vertical transmission from parent(s) to offspring. These transfers allow even distantly related species to share genes, influencing their phenotypes. It is thought that HGT is more prevalent in prokaryotes, but that only about 2% of the



prokaryotic genome may be transferred by this process. Some researchers believe such estimates are premature: the actual importance of HGT to evolutionary processes must be viewed as a work in progress. As the phenomenon is investigated more thoroughly, it may be revealed to be more common. Many scientists believe that HGT and mutation appear to be (especially in prokaryotes) a significant source of genetic variation, which is the raw material for the process of natural selection. These transfers may occur between any two species that share an intimate relationship ([link](#)).

Summary of Mechanisms of Prokaryotic and Eukaryotic HGT			
	Mechanism	Mode of Transmission	Example
Prokaryotes	transformation	DNA uptake	many prokaryotes
	transduction	bacteriophage (virus)	bacteria
	conjugation	pilus	many prokaryotes
	gene transfer agents	phage-like particles	purple non-sulfur bacteria
Eukaryotes	from food organisms	unknown	aphid
	jumping genes	transposons	rice and millet plants

Summary of Mechanisms of Prokaryotic and Eukaryotic HGT			
	Mechanism	Mode of Transmission	Example
	epiphytes/parasites	unknown	yew tree fungi
	from viral infections		

## HGT in Prokaryotes

The mechanism of HGT has been shown to be quite common in the prokaryotic domains of Bacteria and Archaea, significantly changing the way their evolution is viewed. The majority of evolutionary models, such as in the Endosymbiont Theory, propose that eukaryotes descended from multiple prokaryotes, which makes HGT all the more important to understanding the phylogenetic relationships of all extant and extinct species.

The fact that genes are transferred among common bacteria is well known to microbiology students. These gene transfers between species are the major mechanism whereby bacteria acquire resistance to antibiotics. Classically, this type of transfer has been thought to occur by three different mechanisms:

1. Transformation: naked DNA is taken up by a bacteria
2. Transduction: genes are transferred using a virus
3. Conjugation: the use a hollow tube called a pilus to transfer genes between organisms

More recently, a fourth mechanism of gene transfer between prokaryotes has been discovered. Small, virus-like particles called **gene transfer agents (GTAs)** transfer random genomic segments from one species of prokaryote to another. GTAs have been shown to be responsible for genetic changes, sometimes at a very high frequency compared to other evolutionary processes. The first GTA was characterized in 1974 using purple, non-sulfur bacteria.

These GTAs, which are thought to be bacteriophages that lost the ability to reproduce on their own, carry random pieces of DNA from one organism to another. The ability of GTAs to act with high frequency has been demonstrated in controlled studies using marine bacteria. Gene transfer events in marine prokaryotes, either by GTAs or by viruses, have been estimated to be as high as  $10^{13}$  per year in the Mediterranean Sea alone. GTAs and viruses are thought to be efficient HGT vehicles with a major impact on prokaryotic evolution.

As a consequence of this modern DNA analysis, the idea that eukaryotes evolved directly from Archaea has fallen out of favor. While eukaryotes share many features that are absent in bacteria, such as the TATA box (found in the promoter region of many genes), the discovery that some eukaryotic genes were more homologous with bacterial DNA than Archaea DNA made this idea less tenable. Furthermore, the fusion of genomes from Archaea and Bacteria by endosymbiosis has been proposed as the ultimate event in eukaryotic evolution.

## **HGT in Eukaryotes**

Although it is easy to see how prokaryotes exchange genetic material by HGT, it was initially thought that this process was absent in eukaryotes. After all, prokaryotes are but single cells exposed directly to their environment, whereas the sex cells of multicellular organisms are usually sequestered in protected parts of the body. It follows from this idea that the gene transfers between multicellular eukaryotes should be more difficult. Indeed, it is thought that this process is rarer in eukaryotes and has a much smaller evolutionary impact than in prokaryotes. In spite of this fact, HGT between distantly related organisms has been demonstrated in several eukaryotic species, and it is possible that more examples will be discovered in the future.

In plants, gene transfer has been observed in species that cannot cross-pollinate by normal means. Transposons or “jumping genes” have been shown to transfer between rice and millet plant species. Furthermore, fungal species feeding on yew trees, from which the anti-cancer drug TAXOL® is derived from the bark, have acquired the ability to make taxol themselves, a clear example of gene transfer.

In animals, a particularly interesting example of HGT occurs within the aphid species ([\[link\]](#)). Aphids are insects that vary in color based on carotenoid content. Carotenoids are pigments made by a variety of plants, fungi, and microbes, and they serve a variety of functions in animals, who obtain these chemicals from their food. Humans require carotenoids to synthesize vitamin A, and we obtain them by eating orange fruits and vegetables: carrots, apricots, mangoes, and sweet potatoes. On the other hand, aphids have acquired the ability to make the carotenoids on their own. According to DNA analysis, this ability is due to the transfer of fungal genes into the insect by HGT, presumably as the insect consumed fungi for food. A carotenoid enzyme called a desaturase is responsible for the red coloration seen in certain aphids, and it has been further shown that when this gene is inactivated by mutation, the aphids revert back to their more common green color ([\[link\]](#)).



(a)



(b)

(a) Red aphids get their color from red carotenoid pigment. Genes necessary to make this pigment are present in certain fungi, and scientists speculate that aphids acquired these genes through HGT after consuming fungi for food. If genes for making carotenoids are inactivated by mutation, the aphids revert back to (b) their green color. Red coloration makes the aphids a lot more conspicuous to predators, but evidence suggests that red aphids are more resistant to insecticides than green ones. Thus, red aphids may be more fit to survive in some environments than green ones. (credit a: modification of work by Benny Mazur; credit b: modification of work by Mick Talbot)

## Genome Fusion and the Evolution of Eukaryotes

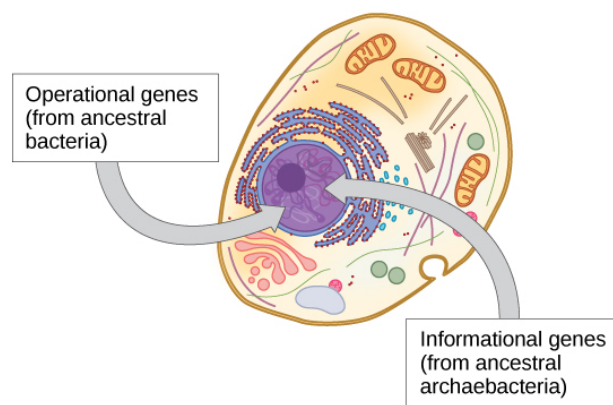
Scientists believe the ultimate in HGT occurs through **genome fusion** between different species of prokaryotes when two symbiotic organisms become endosymbiotic. This occurs when one species is taken inside the cytoplasm of another species, which ultimately results in a genome consisting of genes from both the endosymbiont and the host. This mechanism is an aspect of the Endosymbiont Theory, which is accepted by a majority of biologists as the mechanism whereby eukaryotic cells obtained their mitochondria and chloroplasts. However, the role of endosymbiosis in the development of the nucleus is more controversial. Nuclear and mitochondrial DNA are thought to be of different (separate) evolutionary origin, with the mitochondrial DNA being derived from the circular genomes of bacteria that were engulfed by ancient prokaryotic cells. Mitochondrial DNA can be regarded as the smallest chromosome. Interestingly enough, mitochondrial DNA is inherited only from the mother. The mitochondrial DNA degrades in sperm when the sperm degrades in the fertilized egg or in other instances when the mitochondria located in the flagellum of the sperm fails to enter the egg.

Within the past decade, the process of genome fusion by endosymbiosis has been proposed by James Lake of the UCLA/NASA Astrobiology Institute to be responsible for the evolution of the first eukaryotic cells ([\[link\]a](#)). Using DNA analysis and a new mathematical algorithm called conditioned reconstruction (CR), his laboratory proposed that eukaryotic cells developed from an endosymbiotic gene fusion between two species, one an Archaea and the other a Bacteria. As mentioned, some eukaryotic genes resemble those of Archaea, whereas others resemble those from Bacteria. An endosymbiotic fusion event, such as Lake has proposed, would clearly explain this observation. On the other hand, this work is new and the CR algorithm is relatively unsubstantiated, which causes many scientists to resist this hypothesis.

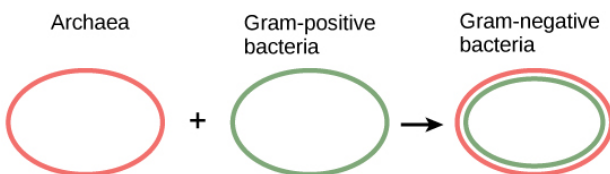
More recent work by Lake ([\[link\]b](#)) proposes that gram-negative bacteria, which are unique within their domain in that they contain two lipid bilayer membranes, indeed resulted from an endosymbiotic fusion of archaeal and bacterial species. The double membrane would be a direct result of the endosymbiosis, with the endosymbiont picking up the second membrane from

the host as it was internalized. This mechanism has also been used to explain the double membranes found in mitochondria and chloroplasts. Lake's work is not without skepticism, and the ideas are still debated within the biological science community. In addition to Lake's hypothesis, there are several other competing theories as to the origin of eukaryotes. How did the eukaryotic nucleus evolve? One theory is that the prokaryotic cells produced an additional membrane that surrounded the bacterial chromosome. Some bacteria have the DNA enclosed by two membranes; however, there is no evidence of a nucleolus or nuclear pores. Other proteobacteria also have membrane-bound chromosomes. If the eukaryotic nucleus evolved this way, we would expect one of the two types of prokaryotes to be more closely related to eukaryotes.

**(a) Genome fusion by endosymbiosis**



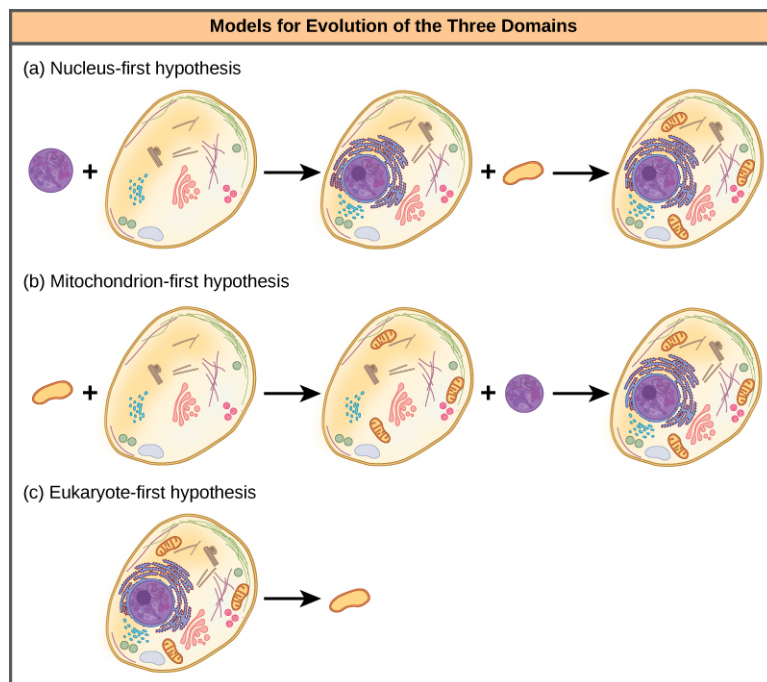
**(b) Endosymbiotic formation of Gram-negative bacteria**



The theory that mitochondria and chloroplasts are endosymbiotic in origin is now widely accepted. More controversial is the proposal that (a) the eukaryotic nucleus resulted from the fusion of archaeal and bacterial genomes, and that (b) Gram-negative

bacteria, which have two membranes, resulted from the fusion of Archaea and Gram-positive bacteria, each of which has a single membrane.

The **nucleus-first** hypothesis proposes that the nucleus evolved in prokaryotes first ([link]a), followed by a later fusion of the new eukaryote with bacteria that became mitochondria. The **mitochondria-first** hypothesis proposes that mitochondria were first established in a prokaryotic host ([link]b), which subsequently acquired a nucleus, by fusion or other mechanisms, to become the first eukaryotic cell. Most interestingly, the **eukaryote-first** hypothesis proposes that prokaryotes actually evolved from eukaryotes by losing genes and complexity ([link]c). All of these hypotheses are testable. Only time and more experimentation will determine which hypothesis is best supported by data.

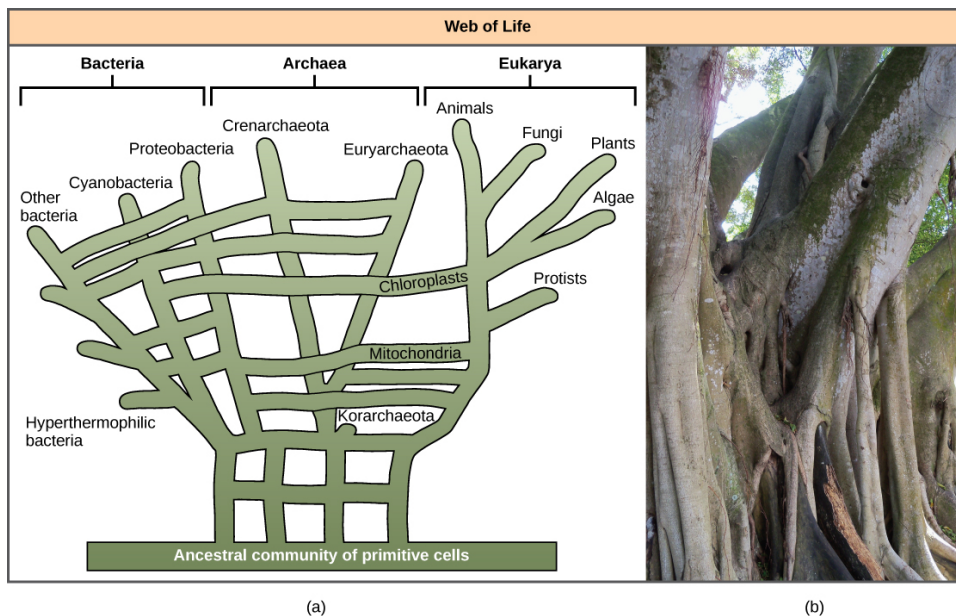


Three alternate hypotheses of eukaryotic and prokaryotic evolution are (a) the nucleus-

first hypothesis, (b) the mitochondrion-first hypothesis, and (c) the eukaryote-first hypothesis.

## Web and Network Models

The recognition of the importance of HGT, especially in the evolution of prokaryotes, has caused some to propose abandoning the classic “tree of life” model. In 1999, W. Ford Doolittle proposed a phylogenetic model that resembles a web or a network more than a tree. The hypothesis is that eukaryotes evolved not from a single prokaryotic ancestor, but from a pool of many species that were sharing genes by HGT mechanisms. As shown in [\[link\]a](#), some individual prokaryotes were responsible for transferring the bacteria that caused mitochondrial development to the new eukaryotes, whereas other species transferred the bacteria that gave rise to chloroplasts. This model is often called the “**web of life**.” In an effort to save the tree analogy, some have proposed using the *Ficus* tree ([\[link\]b](#)) with its multiple trunks as a phylogenetic to represent a diminished evolutionary role for HGT.



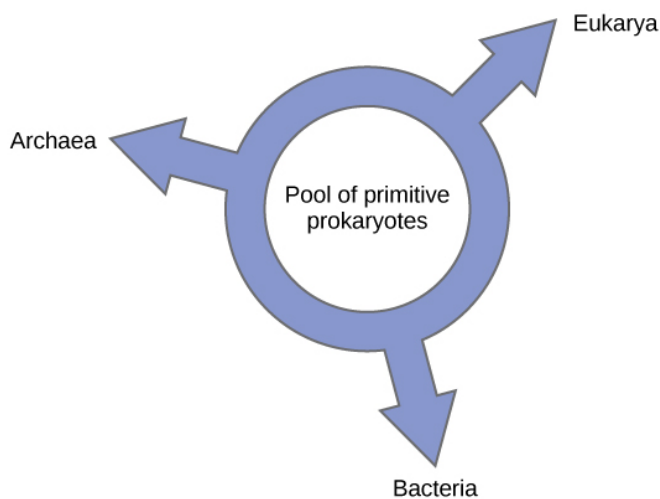
In the (a) phylogenetic model proposed by W. Ford



Doolittle, the “tree of life” arose from a community of ancestral cells, has multiple trunks, and has connections between branches where horizontal gene transfer has occurred. Visually, this concept is better represented by (b) the multi-trunked **Ficus** than by the single trunk of the oak similar to the tree drawn by Darwin [\[link\]](#).  
(credit b: modification of work by "psyberartist"/Flickr)

## Ring of Life Models

Others have proposed abandoning any tree-like model of phylogeny in favor of a ring structure, the so-called “**ring of life**” ([\[link\]](#)); a phylogenetic model where all three domains of life evolved from a pool of primitive prokaryotes. Lake, again using the conditioned reconstruction algorithm, proposes a ring-like model in which species of all three domains—Archaea, Bacteria, and Eukarya—evolved from a single pool of gene-swapping prokaryotes. His laboratory proposes that this structure is the best fit for data from extensive DNA analyses performed in his laboratory, and that the ring model is the only one that adequately takes HGT and genomic fusion into account. However, other phylogeneticists remain highly skeptical of this model.



According to the “ring of life”

phylogenetic model, the three domains of life evolved from a pool of primitive prokaryotes.

In summary, the “tree of life” model proposed by Darwin must be modified to include HGT. Does this mean abandoning the tree model completely? Even Lake argues that all attempts should be made to discover some modification of the tree model to allow it to accurately fit his data, and only the inability to do so will sway people toward his ring proposal.

This doesn’t mean a tree, web, or a ring will correlate completely to an accurate description of phylogenetic relationships of life. A consequence of the new thinking about phylogenetic models is the idea that Darwin’s original conception of the phylogenetic tree is too simple, but made sense based on what was known at the time. However, the search for a more useful model moves on: each model serving as hypotheses to be tested with the possibility of developing new models. This is how science advances. These models are used as visualizations to help construct hypothetical evolutionary relationships and understand the massive amount of data being analyzed.

## **Section Summary**

The phylogenetic tree, first used by Darwin, is the classic “tree of life” model describing phylogenetic relationships among species, and the most common model used today. New ideas about HGT and genome fusion have caused some to suggest revising the model to resemble webs or rings.

## **Review Questions**

### **Exercise:**

#### **Problem:**

The transfer of genes by a mechanism not involving asexual reproduction is called:

- a. meiosis

- b. web of life
- c. horizontal gene transfer
- d. gene fusion

---

**Solution:**

C

**Exercise:**

**Problem:**

Particles that transfer genetic material from one species to another, especially in marine prokaryotes:

- a. horizontal gene transfer
- b. lateral gene transfer
- c. genome fusion device
- d. gene transfer agents

---

**Solution:**

D

**Exercise:**

**Problem:** What does the trunk of the classic phylogenetic tree represent?

- a. single common ancestor
- b. pool of ancestral organisms
- c. new species
- d. old species

---

**Solution:**

A

**Exercise:**

**Problem:**

Which phylogenetic model proposes that all three domains of life evolved from a pool of primitive prokaryotes?

- a. tree of life
- b. web of life
- c. ring of life
- d. network model

---

**Solution:**

C

**Free Response****Exercise:****Problem:**

Compare three different ways that eukaryotic cells may have evolved.

---

**Solution:**

Some hypotheses propose that mitochondria were acquired first, followed by the development of the nucleus. Others propose that the nucleus evolved first and that this new eukaryotic cell later acquired the mitochondria. Still others hypothesize that prokaryotes descended from eukaryotes by the loss of genes and complexity.

**Exercise:**

**Problem:** Describe how aphids acquired the ability to change color.

---

**Solution:**

Aphids have acquired the ability to make the carotenoids on their own. DNA analysis has demonstrated that this ability is due to the transfer of

fungus genes into the insect by HGT, presumably as the insect consumed fungi for food.

## **Glossary**

eukaryote-first hypothesis

proposal that prokaryotes evolved from eukaryotes

gene transfer agent (GTA)

bacteriophage-like particle that transfers random genomic segments from one species of prokaryote to another

genome fusion

fusion of two prokaryotic genomes, presumably by endosymbiosis

horizontal gene transfer (HGT)

(also, lateral gene transfer) transfer of genes between unrelated species

mitochondria-first hypothesis

proposal that prokaryotes acquired a mitochondrion first, followed by nuclear development

nucleus-first hypothesis

proposal that prokaryotes acquired a nucleus first, and then the mitochondrion

ring of life

phylogenetic model where all three domains of life evolved from a pool of primitive prokaryotes

web of life

phylogenetic model that attempts to incorporate the effects of horizontal gene transfer on evolution

# The Periodic Table of Elements

Periodic Table of the Elements																		18																	
1		2												13	14	15	16	17																	
1	H 1.01 Hydrogen													5	B 10.81 Boron	6	C 12.11 Carbon	7	N 14.01 Nitrogen	8	O 15.99 Oxygen	9	F 18.99 Fluorine	10	Ne 20.18 Neon										
3	Li 6.94 Lithium	4	Be 9.01 Beryllium											13	Al 26.98 Aluminum	14	Si 28.09 Silicon	15	P 30.97 Phosphorus	16	S 32.07 Sulfur	17	Cl 35.45 Chlorine	18	Ar 39.95 Argon										
11	Na 22.99 Sodium	12	Mg 24.31 Magnesium											31	Ga 69.72 Gallium	32	Ge 72.64 Germanium	33	As 74.92 Arsenic	34	Se 78.96 Selenium	35	Br 79.90 Bromine	36	Kr 83.79 Krypton										
19	K 39.09 Potassium	20	Ca 40.08 Calcium	21	Sc 44.96 Scandium	22	Ti 47.87 Titanium	23	V 50.94 Vanadium	24	Cr 51.99 Chromium	25	Mn 54.94 Manganese	26	Fe 55.85 Iron	27	Co 58.93 Cobalt	28	Ni 58.69 Nickel	29	Cu 63.55 Copper	30	Zn 65.41 Zinc	31	Ga 69.72 Gallium	32	Ge 72.64 Germanium	33	As 74.92 Arsenic	34	Se 78.96 Selenium	35	Br 79.90 Bromine	36	Kr 83.79 Krypton
37	Rb 85.47 Rubidium	38	Sr 87.62 Strontium	39	Y 88.91 Yttrium	40	Zr 91.22 Zirconium	41	Nb 92.91 Niobium	42	Mo 95.94 Molybdenum	43	Tc [98] Technetium	44	Ru 101.1 Ruthenium	45	Rh 102.9 Rhodium	46	Pd 106.4 Palladium	47	Ag 107.9 Silver	48	Cd 112.4 Cadmium	49	In 114.8 Indium	50	Sn 118.7 Tin	51	Sb 121.8 Antimony	52	Te 127.6 Tellurium	53	I 126.9 Iodine	54	Xe 131.3 Xenon
55	Cs 132.9 Cesium	56	Ba 137.3 Barium	57-71	La-Lu [La-Lu]	72	Hf 178.5 Hafnium	73	Ta 180.9 Tantalum	74	W 183.8 Tungsten	75	Re 186.2 Rhenium	76	Os 190.2 Osmium	77	Ir 192.2 Iridium	78	Pt 195.1 Platinum	79	Au 196.9 Gold	80	Hg 200.6 Mercury	81	Tl 204.4 Thallium	82	Pb 207.2 Lead	83	Bi 208.9 Bismuth	84	Po [209] Polonium	85	At [210] Astatine	86	Rn [222] Radon
87	Fr [223] Francium	88	Ra [226] Radium	89-103	Ac-Lr [Ac-Lr]	104	Rf [261] Rutherfordium	105	Db [262] Dubnium	106	Sg [266] Seaborgium	107	Bh [264] Bohrium	108	Hs [277] Hassium	109	Mt [268] Meitnerium	110	Ds [269] Darmstadtium	111	Rg [272] Roentgenium	112	Cn [285] Copernicium	113	Uut [284] Ununtrium	114	Fl [289] Flerovium	115	Uup [288] Ununpentium	116	Lv [293] Livermorium	117	Uus [294] Ununseptium	118	Uuo [294] Ununoctium
				57	La 138.9 Lanthanum	58	Ce 140.1 Cerium	59	Pr 140.9 Praseodymium	60	Nd 144.2 Neodymium	61	Pm [145] Promethium	62	Sm 150.4 Samarium	63	Eu 151.9 Europium	64	Gd 157.3 Gadolinium	65	Tb 158.9 Terbium	66	Dy 162.5 Dysprosium	67	Ho 164.9 Holmium	68	Er 167.3 Erbium	69	Tm 168.9 Thulium	70	Yb 173.1 Ytterbium	71	Lu 174.9 Lutetium		
				89	Ac [227] Actinium	90	Th 232.0 Thorium	91	Pa 231.0 Protactinium	92	U 238.0 Uranium	93	Np [237] Neptunium	94	Pu [244] Plutonium	95	Am [243] Americium	96	Cm [247] Curium	97	Bk [247] Berkelium	98	Cf [251] Californium	99	Es [252] Einsteinium	100	Fm [257] Fermium	101	Md [258] Mendelevium	102	No [259] Nobelium	103	Lr [262] Lawrencium		

Atomic Number → 1

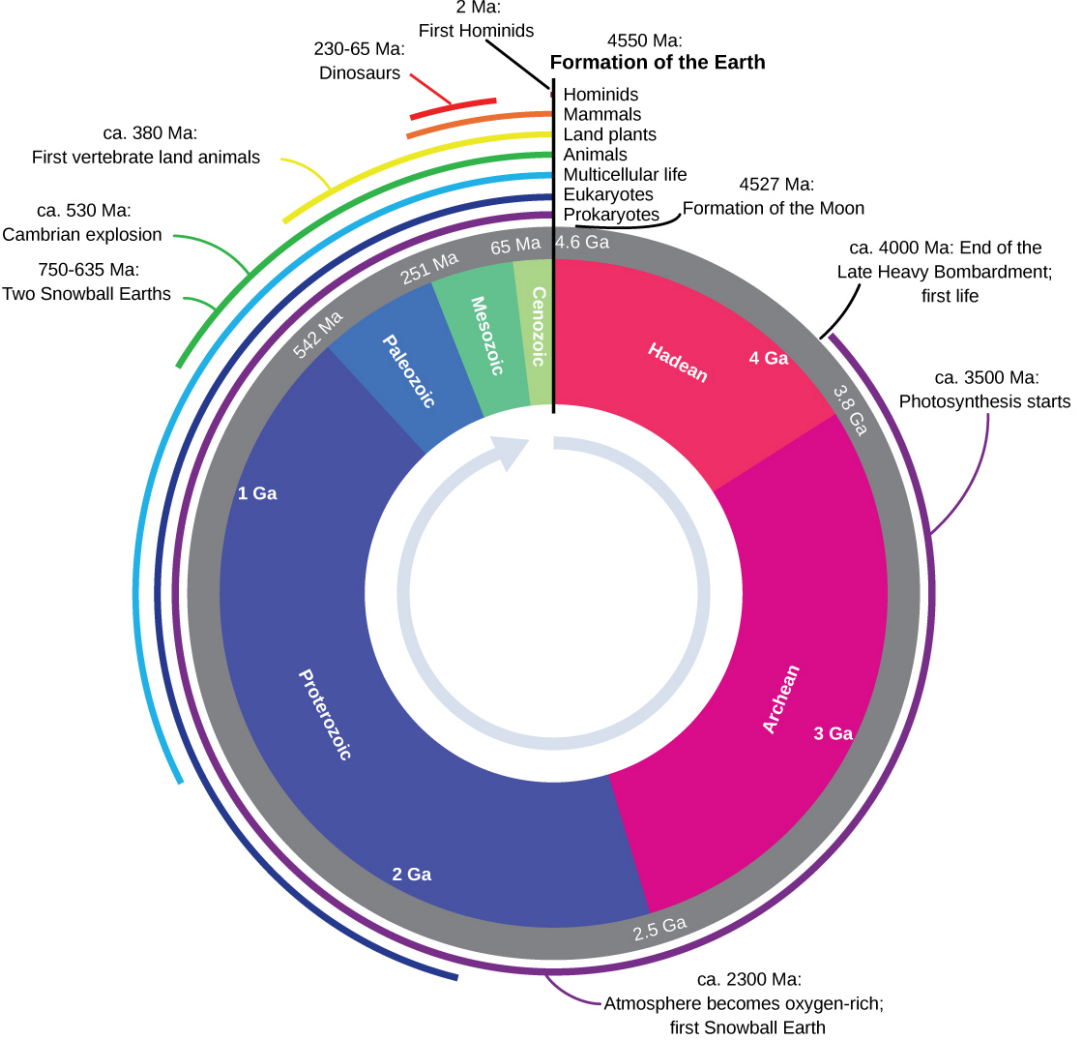
Symbol → H

Relative Atomic Mass → 1.01

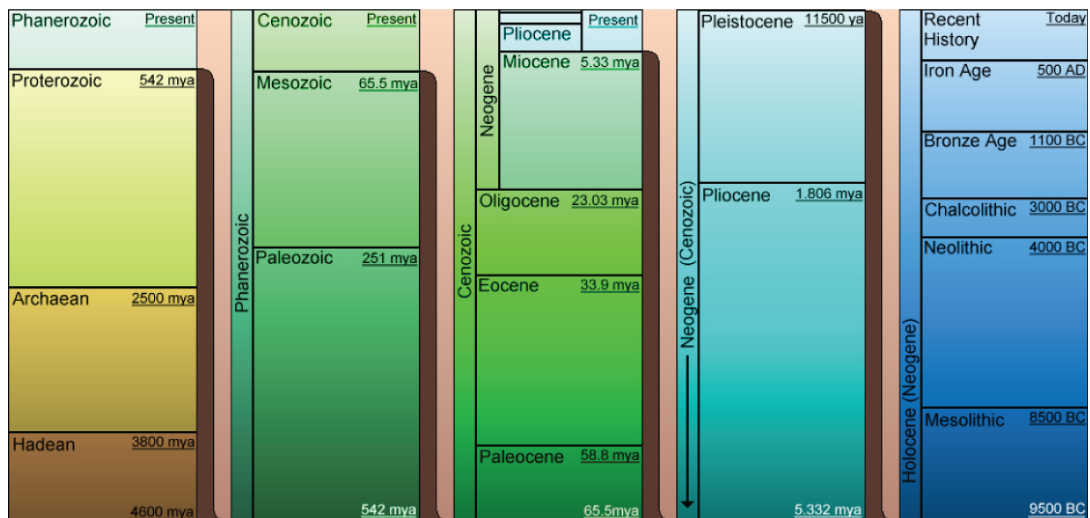
Name → Hydrogen

Color Code	
<div></div> Other non-metals	<div></div> Noble gases
<div></div> Alkali metals	<div></div> Lanthanides
<div></div> Transition metals	<div></div> Actinides
<div></div> Other metals	<div></div> Unknown chemical properties
<div></div> Alkaline earth metals	
<div></div> Halogens	

# Geological Time



Geological Time Clock



Geological Time Chart  
(credit: Richard S. Murphy, Jr.)



Measurements and the Metric System

Measurements and the Metric System				
Measurement	Unit	Abbreviation	Metric Equivalent	Approximate Standard Equivalent
Length	nanometer	nm	1 nm = $10^{-9}$ m	<ul style="list-style-type: none"><li>• 1 mm = 0.039 inch</li><li>• 1 cm = 0.394 inch</li><li>• 1 m = 39.37 inches</li><li>• 1 m = 3.28 feet</li><li>• 1 m = 1.093 yards</li><li>• 1 km = 0.621 miles</li></ul>
	micrometer	μm	1 μm = $10^{-6}$ m	
	millimeter	mm	1 mm = 0.001 m	
	centimeter	cm	1 cm = 0.01 m	
	meter	m	<ul style="list-style-type: none"><li>• 1 m = 100 cm</li><li>• 1 m = 1000 mm</li></ul>	
	kilometer	km	1 km = 1000 m	
Mass	microgram	μg	1 μg = $10^{-6}$ g	<ul style="list-style-type: none"><li>• 1 g = 0.035 ounce</li><li>• 1 kg = 2.205 pounds</li></ul>
	milligram	mg	1 mg = $10^{-3}$ g	
	gram	g	1 g = 1000 mg	

Measurements and the Metric System				
Measurement	Unit	Abbreviation	Metric Equivalent	Approximate Standard Equivalent
	kilogram	kg	1 kg = 1000 g	
Volume	microliter	μl	1 μl = $10^{-6}$ l	<ul style="list-style-type: none"> <li>• 1 ml = 0.034 fluid ounce</li> <li>• 1 l = 1.057 quarts</li> <li>• 1 kl = 264.172 gallons</li> </ul>
	milliliter	ml	1 ml = $10^{-3}$ l	
	liter	l	1 l = 1000 ml	
	kiloliter	kl	1 kl = 1000 l	
Area	square centimeter	cm <sup>2</sup>	1 cm <sup>2</sup> = 100 mm <sup>2</sup>	<ul style="list-style-type: none"> <li>• 1 cm<sup>2</sup> = 0.155 square inch</li> <li>• 1 m<sup>2</sup> = 10.764 square feet</li> <li>• 1 m<sup>2</sup> = 1.196 square yards</li> <li>• 1 ha = 2.471 acres</li> </ul>
	square meter	m <sup>2</sup>	1 m <sup>2</sup> = 10,000 cm <sup>2</sup>	
	hectare	ha	1 ha = 10,000 m <sup>2</sup>	
Temperature	Celsius	°C	—	1 °C = $\frac{5}{9} \times$ (°F – 32)